The complex network topology of trade in a globalized world

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m
(Dr. rer. nat.)
im Fach Physik
Spezialisierung: Theoretische Physik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Dipl.-Phys. Julian Maluck

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:
Prof. Dr. Elmar Kulke

Gutachter:

1. Prof. Dr. Dr. h.c. mult. Jürgen Kurths
2. Prof. Dr. Anders Levermann
3. PD Dr. Jens Christian Claussen

Tag der mündlichen Prüfung: 3. Mai 2018

Abstract

The economy and, more specifically, the trade of goods and services can arguably be regarded as a complex system. Trade is omnipresently affecting our everyday life as an economic agent, whether as a consumer or as a producer, during leisure time or in work life. Therefore, the organization of trade and its patterns and structures have always had far reaching implications among social, political and economic dimensions. Trade is in its essence the outcome of many interactions among different agents. Thus, the underlying dynamics of trade systems are often a priori unknown and concepts from complex system theory provide useful tools to discover new patterns and to develop new hypotheses on the mechanisms of the system. Complex networks offer a particularly useful approach to trade systems, as trade flows between economic entities can be intuitively and meaningfully represented as nodes and links in a network.

In this thesis, we extend specific methods of complex networks with a focus on the relations between different subnetworks to investigate the network topology of trade on both the global and national scale. On an aggregation level considering individual industries as nodes, we obtain new insights about the topological structure of the international trade network by introducing new network measures that characterize the roles of nodes in subnetworks from a network of networks perspective. Furthermore, we extend existing methods to weighted and directed networks that allow the identification of events that result in large reorganization processes of the network's link structure. In the context of trade, these events define crises and economic shocks.

During the process of globalization bilateral trade agreements have received rising attention among policy makers and have been negotiated at an increasing pace. Here, we develop a framework to analyze and quantify impacts of these agreements on the involved economies. Thus, we evaluate the interconnectedness between two subnetworks by considering path probabilities that are deduced from the trade volumes. We introduce a method to categorize the evolution of a measure on short time scales that take both trend and level considerations into account.

A further question that we address in this thesis is to what extent trade can be regarded as a mediator of demand and supply spillovers to other industries. To this end, we present empirical methods to compare the topologies of the international trade network with a functional network of economic performance. Finally, we look into trade networks at the scale of individual business firms and describe the role of nodes with a focus on 3-node motifs. With simulations of an evolving network model we learn about the underlying mechanisms in the formation process of motifs in a scale-free network topology. We illustrate how this methodology provides new insights into the preferences of Japanese businesses in choosing their trade relationships. Although all new methods and measures introduced in this thesis are motivated by questions in the context of trade, the methodological concepts are widely applicable to complex networks of other research disciplines.

Zusammenfassung

Die Wirtschaft und, genauer gesagt, der Handel von Waren und Dienstleistungen kann als ein komplexes System aufgefasst werden. Der Warenhandel ist allgegenwärtiger Bestandteil in unserem Alltag als ökonomisch handelndes Individuum, sei es in der Rolle als Konsument oder als Produzent, sei es in unserer Freizeit oder im Berufsleben. Dadurch brachte die Organisation von Handelsstrukturen seit je her weitreichende soziale, politische sowie ökonomische Implikationen mit sich. Grundsätzlich kann Handel als Resultat von vielen Interaktionen verschiedener ökonomischer Handlungen aufgegasst werden. Die zugrundeliegenden Dynamiken von Handelssystemen sind dabei a priori unbekannt. Konzepte aus der Theorie komplexer Systeme bieten dabei nützliche Werkzeuge, um neue Muster zu entdecken, sowie neue Hypothesen zu den Vorgängen innerhalb der Handelssysteme zu entwickeln. Einen ausgesprochen nützlichen Ansatz stellen dabei komplexe Netzwerke dar, da die Handelsströme zwischen ökonomischen Einheiten intuitiv und sinnvoll als Knoten und Verbindungen im Netzwerk dargestellt werden können.

In dieser Arbeit erweitern wir spezielle Methoden auf komplexen Netzwerken, um die Netzwerk-Topologie des Handels sowohl auf globaler als auch auf nationaler Ebene zu untersuchen. Auf dem Level einzelner Industriesektoren als Knoten erhalten wir neue Einblicke in die topologische Struktur des internationalen Handelsnetzwerks. Dies gelingt, indem wir neue Netzwerkmaße einführen, welche die Funktion von Knoten in Subgraphen unter dem Blickwinkel, dass das Gesamtsystem durch ein Netzwerk aus mehreren Subnetzwerken dargestellt wird, beschreiben. Weiterhin erweitern wir existierende Methoden für die Anwendung auf gewichtete und gerichtete Netzwerke. Damit können wir Ereignisse identifizieren, die viele Reorganisationsprozesse in der Linkstruktur des Netzwerks verursachen. Im Handelskontext definieren diese Ereignisse Krisen und ökonomische Schocks.

Im Zuge der Globalisierung gewinnen bilaterale Handelsabkommen bei Entscheidungsträgern an Aufmerksamkeit und werden in zunehmender Zahl verhandelt. Wir entwickeln in dieser Arbeit einen neuen Ansatz, um die Auswirkungen dieser Abkommen auf die beteiligten Ökonomien zu analysieren und zu quantifizieren. Dazu evaluieren wir den Verflechtungsgrad zwischen zwei Subnetzwerken, indem wir die Wahrscheinlichkeiten des Folgens von Netzwerkpfaden berücksichtigen. Diese Wahrscheinlichkeiten werden dabei aus den Handelsvolumina abgeleitet. Wir führen eine weitere Methode ein, um den Verlauf eines Maßes auf kurzen Zeitskalen unter Berücksichtigung von Trends und Niveauveränderungen zu kategorisieren.

Diese Arbeit beschäftigt sich weiterhin mit der Fragestellung, in welchem Maß Handel als ein Übermittler von Nachfrage- und Angebotsveränderungen auf andere Industrien angesehen werden kann. Dazu präsentieren wir empirische Methoden, um die Topologien des internationalen Handelsnetzwerks mit der eines Netzwerks, welches die funktionellen Zusammenhänge der ökonomischen Leistungsfähigkeiten beschreibt, zu vergleichen. Schließlich betrachten wir Handelsnetzwerke auf der Ebene von einzelnen Firmen und beschreiben die Funktion von einzelnen Knoten, insbesondere deren Rolle innerhalb von 3er-Motiven. Mithilfe von Simulationen eines evolvierenden Netzwerkmodells erhalten wir

neue Einblicke über die zugrundeliegenden Mechanismen im Formationsprozess von Motiven in skalenfreien Netzwerken. Wir zeigen anschaulich, wie diese Methodik neue Erkenntnisse über Präferenzen von japanischen Firmen liefert, wenn diese nach Handelsbeziehungen suchen. Obwohl alle Methoden und Maße, die wir im Zuge dieser Arbeit einführen, aus Fragen im Kontext des Handels motiviert sind, sind die methodischen Konzepte auf komplexe Netzwerke in anderen Forschungsrichtungen anwendbar.

List of publications

This dissertation is partly based on the following publications. The identifiers given below are cited in the text to highlight passages that are connected to one or more of these papers.

Papers

P1 **J. Maluck** and R. V. Donner. "A Network of Networks Perspective on Global Trade". In: *PLoS ONE* 10(7) (2015), e0133310

P2 **J. Maluck** and R. V. Donner. "Distributions of positive correlations in sectoral value added growth in the global economic network". In: *European Physical Journal B* 90(2) (2017), 26

P3 **J. Maluck**, R. V. Donner, H. Takayasu and M. Takayasu. "Motif formation and industry specific topologies in the Japanese Business Firm Network". In: *Journal of Statistical Mechanics: Theory and Experiment* 5 (2017), 053404

P4 **J. Maluck**, N. Glanemann and R. V. Donner, "Bilateral trade agreements and the interconnectedness of global trade - The different roles played by the United States and China". *(submitted)*

Acknowledgements

This work would not have been possible without the great supervision of Reik Donner and Prof. Jürgen Kurths. I am very grateful for their unlimited support, their confidence in me and their affirmations and help for all the plans I came up with during my time at PIK. I want to express an extra thanks to Reik for always keeping me in the right scientific track, for teaching me the art of network science in person and for always finding the appropriate funds somewhere. I am very indebted to Profs. Misako and Hideki Takayasu for the very successful collaboration and for providing me the unique and exciting opportunity to experience the life as a researcher in Japan.

Next, I would like to mention the CoSy Team: thanks to Jasper and Chiranjit I received a constant supply of the best tea choices as well as Indian delicacies. Thank you Marc, Jaqueline and Jonatan for steering together in the same boat and helping me out whenever you could. To look beyond CoSy, I want to thank Paul very much for all the interesting stuff he told me. One particular big thanks goes to Deniz, for showing me how scientific life really works, for providing me guidance in so many things and simply for all the great times. I am also grateful to Carsten for keeping me motivated and for his key improvements in my application for the Japanese scholarship. Talking of Japan, without Takumi Sueshige and Maki Torii I would certainly have not been able to live a life there and still struggle to register for anything. More thanks go to Leonie and Alex for showing and guiding me through trade networks at the very crucial time of the beginning of my research. And thank you Nicole, for being a great and inspiring collaborator.

Last but not least I would like to thank the IT-Crew at PIK for providing a most reliable computing system and being ready for any support at any time. Furthermore, I am very grateful to the BMBF, the JSPS and the DAAD. These are the institutions that made this work possible by supporting it financially.

Contents

List of Figures

List of Tables

List of frequently used mathematical symbols

v_i	a node, indexed by i
(i, j)	a link between nodes v_i and v_j
$\mathcal{V}$	set of nodes / vertices
$\mathcal{E}$	set of links / edges
$\mathcal{G}(\mathcal{V}, \mathcal{E})$	a network with the set of nodes $\mathcal{V}$ and the set of links $\mathcal{E}$
m	number of links in the network
N	number of nodes in the network
$\mathbf{A}$, $(\mathbf{A}')$	Adjacency matrix of a directed (undirected) network
a_{ij}	one entry of $\mathbf{A}$
$\mathbf{W}$, $(\mathbf{W}')$	Weight matrix of a directed (undirected) network
w_{ij}	one entry of $\mathbf{W}$
$\mathbf{T}$	intermediate trade block in the input-output table
$\mathbf{FD}$	final demand block in the input-output table
$\mathbf{VA}$	value added block in the input-output table
k_i	degree of node v_i
s_i	strength of node v_i
p_{ij}^{out}	output dependency of node v_i on v_j
p_{ij}^{in}	input dependency of node v_j on v_i
$\mathbf{P}_\bullet$	weight matrix of the flow network with $(\mathbf{P}_\bullet)_{ij} = p_{ij}^\bullet$ and $\bullet \in \{in, out\}$
C_i	local clustering coefficient of node v_i
$C_i^\bullet$	local clustering coefficient for a specific pattern $\bullet \in \{cyc, mid, in, out, all\}$ in directed networks
σ_{jk}	number of shortest paths from node v_j to v_k
b_i	betweenness of node v_i
B^{pq}	cross-betweenness fraction of shortest paths from subnetwork p to q that pass through a third subnetwork
ρ	link density of the network
r	reciprocity of the network
$H(\mathcal{G}_x, \mathcal{G}_y)$	Hamming distance between the networks $\mathcal{G}_x$ and $\mathcal{G}_y$

Φ	average link weight of the network
l_b	blinking links in the Hamming distance
Δw	difference in strength of links
$\mathcal{C}$	partition of a network into subnetworks
$Q(\mathcal{C})$	modularity of a network partition $\mathcal{C}$
$TI(\mathcal{G}'_x, \mathcal{G}'_y)$	trade interconnectedness between the subnetworks $\mathcal{G}'_x$ and $\mathcal{G}'_y$
$\Pi(\mathcal{G}'_x, \mathcal{G}'_y)$	BTA impact index between the subnetworks $\mathcal{G}'_x$ and $\mathcal{G}'_y$
VA_i	value added of node v_i
νa_i	value added growth of node v_i
g_i	corrected value added growth of node v_i
$M(x)$	estimated median of random variable x
$QR^{\pm}(x)$	estimated upper $(+)$ and lower $(-)$ quartile from a sample of random variable x
Corr_{ij}	statistical association between two entities i and j
$\mathrm{E}(x)$	estimated mean from a sample of random variable x
$\mathrm{Var}(x)$	estimated variance from a sample of random variable x
$\mathrm{ACF}(x)$	estimated autocorrelation from a time series of random variable x
KS	Kolmogorov-Smirnov statistic between two probability distributions
JS	Jenson-Shannon divergence between two probability distributions
c_{ij}^{in}	common neighborhood input dependency between v_i and v_j
c_{ij}^{out}	common neighborhood output dependency between v_i and v_j
$\mathcal{VIN}_{ij}$	common international neighborhood of v_i and v_j
cn_{ij}^{in}	common international neighborhood input dependency between v_i and v_j
cn_{ij}^{out}	common international neighborhood output dependency between v_i and v_j
μ	motif of a specific pattern μ
R	role of the node in a motif pattern
$\Delta P_\mu(\mathcal{G}_x, \mathcal{G}_y)$	penalty value of motif pattern μ between two networks $\mathcal{G}_x$ and $\mathcal{G}_y$

Chapter 1.

Introduction

Trade has always played an essential role within human cultures, societies and economies. It forms a key pillar for a society's wealth and for technological progress: No single person is able to grow their own food, build all the desired infrastructure on their own and produce the goods that they wish to consume alone. Trading activities allow for the division of labor and for people to specialize in certain tasks. The concept of specialization is crucial to humanity and builds a prerequisite for the improvement of mankind's knowledge as well as for the improvement of the efficiency and of the technology in any working process. In today's global economy, trade occurs at different scales in different magnitudes. Individual people trade among each other, but also businesses and larger corporations, and even whole industries and nations. Thus, trade has been connecting the world with all its cultural and political influences in the process of globalization. Although trade plays a seemingly very intuitive role among individual people in our everyday life and in our daily consumption, there is continuous debate on the organization of trade. The impacts of trade on a society and on the people's lives are far reaching and there is a lot of discussion about the attribution of winners and losers of specific forms of trade [1–3].

The debates on trade are fueled by the fact that the economy is a complex system that has yet to be understood in its entirety. In traditional economic theory, a popular approach includes the development of models to investigate the economic question of interest [4, 5]. Another branch of research can be attributed to econometric data analyses that are often based on regression techniques [6]. The general philosophy in both cases is to reduce complexity by investigating specific economic mechanisms in isolation from the rest of the economic system. Within this isolated environment, mostly linear models are then utilized to deduce predictions given a set of assumptions. These assumptions often postulate the behavioral patterns of the economic entities or people in the model. The concept of complexity reduction by isolating subsystems is also a fundamental approach that is pursued in the natural sciences and in physics. When we investigate the law of gravity by throwing a stone, we are less concerned about the rotation of the earth or the fact that a train is passing nearby. Physical models are tested by experiments. These have to be designed in such a way that they resemble the isolated subsystem under investigation. The mechanisms that can significantly influence the outcome of the experiments should be consistently describable by a model. In economics, however, these experiments and data on isolated subsystems that would allow for tests of the economic models are in general not

available. The economy is in its essence the outcome of complex interactions between people that build organizations and form economic entities such as corporations or nations. The mechanisms that influence certain subsystems of the economy are therefore hard to identify. For example, while it seems certainly reasonable to assume that the municipal policies of Berlin have a negligible impact on the trade relationship between Japan and China, many factors that range from natural disasters to political or economic sentiment of the people can potentially have a significant impact. Thus, testing economic models with data analyses has in general been a challenging and cumbersome task that often allows for various different interpretations of the results.

Concepts from complex system theory offer promising alternative approaches to investigate patterns within economic systems. In particular, complex network theory provides useful tools to analyze the topology of trade relations among the economic entities. Networks, or graphs as they are also called, describe the interconnections between a defined set of entities that are represented as nodes [7, 8]. Conceptually the theory builds on mathematical graph theory which is said to originate from Euler's question on how to cross all bridges in Königsberg once and only once [9]. In modern times, real-world systems of many research disciplines have been analyzed and interpreted as a complex network, ranging from social and economic sciences [10, 11], biology [12, 13], computer science [14], neuroscience [15] to earth sciences [16]. Network analyses allow for the identification of structures and patterns at both large and small scale that emerge in the system under study. At the small scale, local properties of the nodes allow the identification of specific functional roles of individual nodes in the network. On the other hand, global network properties provide insights to characteristic properties of the system as a whole. If the mechanisms of the system are a priori unknown, the understanding of its network topology and evolution plays often a key role in finding and formulating these mechanisms. In the context of economics, for example, it has been found that the core-periphery structure of a network of industry products has implications on a country's opportunity for economic growth [17]. In biological sciences, the scale-free property in protein-protein interaction networks provides evidence that the connections of new proteins to the existing network is governed by preferential attachment [18]. The topological properties of a network's structure can also provide insights to dynamical processes on the networks. For example, with the knowledge of the link structure in social networks, the spreading of diseases [19] and computer viruses [20] can be efficiently modeled.

As the diversity of the examples above illustrate, another advantage of complex networks is its universal applicability. Networks can represent systems with any kind of node interactions that can be governed by any kind of dynamics, i.e. linear or non-linear. The complex network approach can only be reasonable, however, if the network and the measures therein can be meaningfully interpreted and understood. In the context of trade, complex networks offer an intuitive visualization of trade systems with the nodes representing the economic entities and the links representing the trade flow in form of exchanged goods or services. Previous studies have discovered interesting patterns in trade networks [21–23]. Next to observations of previously

known developments in trade systems that served as a proof of concept of the methodology, new patterns that were a priori unknown have been identified as well in these studies. For example, the scale-free topology of global trade on the level of countries suggests that the evolution of the network is guided by self-organization processes [21]. Furthermore, it has been shown that the GDP of countries serve as key determinants for the probability that the countries engage in trade relations with each other [22]. Other work suggests that the complex network approach is suitable to quantify a country's importance and influence in the network of global trade [24].

In this thesis, we want to obtain a more detailed picture on trade systems by investigating trade relations at the scale of individual industries and businesses. Therefore, we extend existing methods of complex network theory and adapt it to the research questions of interest in order to allow for meaningful interpretations in the specific context. There are four main questions that we address in this thesis: First, we investigate how globalization has affected the network of global trade at the scale of industry sectors in the last decades. For that purpose, we consider the trade network as a network of networks and evaluate its evolution. During the same time period, the world has witnessed a surge in implementations of bilateral trade agreements, receiving rising attention among policy makers [25]. Thus, in the second focus of this thesis we want to understand the impact of these agreements on the structure of global trade and on the interconnectedness between the involved economies. In an ever more interconnected global economy, the question arises how the economic system is affected with regard to its vulnerability to specific shocks of certain industries [26–28]. To empirically find possible factors that affect the spreading mechanisms of shocks, we introduce a framework to study the interrelation between the trade of industries within the network and their economic performance. Finally, we investigate the formation of trade relationships between individual firms. In particular, we are interested in the preferences of firms in building new business relations and in their role within their neighborhood within the network.

We begin with a description of the concepts that are utilized throughout this thesis in chapter 2. Next to the presentation of the relevant measures and methods from complex network theory, we discuss here their interpretation and introduce some new meaningful measures in the context of trade networks. Most of the data analyzed in this thesis is based on concepts of economic input-output analysis that we present as well. In the following chapter 3, we investigate the network topology of global trade on the scale of industries and identify meaningful subnetworks to characterize their interactions among each other. Conceptually, we interpret trade as a network of networks and extend the methodology and the measures to this context. To quantify temporal changes in the link structure of the trade network, we generalize the concept of the Hamming distance [29] to weighted networks. Having obtained a more general overview and understanding of the global trade network and its time evolution, we study the question of how bilateral trade agreements have affected the trade relationships between the involved countries in chapter 4. For this purpose, we develop a methodology to quantify the interconnectedness between two economies and analyze the time evolution of this measure. We provide a global overview of the

impacts of bilateral trade agreements and analyze the trade agreement profiles of two of the world's leading economic powers in detail, namely the profiles of the USA and China. It turns out that the bilateral trade agreements of these two nations show pronounced differences in their impact.

It is often hypothesized that trade serves as a mediator of shocks of supply and demand [30]. In chapter 5 we address this question empirically and assess if trade relations between industries and the resulting network topology provide insights into correlations in the industries' economic performance. From a conceptual perspective, we construct a functional network of the correlation structure of economic performance and compare its topology with the corresponding characteristics of the trade network.

Finally, we investigate the topology of trade at the scale of business firms within the national economy of Japan in chapter 6. Here, we utilize a model approach and compare the statistical distribution of motifs, i.e. small induced subnetworks, in the model with the corresponding distribution in the real-world network. Thus, we obtain new conceptual knowledge about the formation processes of motifs in a scale-free network topology. We conclude the thesis with a summary and discussion in chapter 7.

Chapter 2.

Theoretical framework

The networks that are subject to this thesis describe the trade of goods and services. They represent an aggregation of all the supply chains of the individual products from the entire product range of the economy. Complex network theory provides a suitable toolbox to analyze and better understand the underlying structures and topological properties of the flow of goods and money within these trade networks. In this chapter, we discuss the utilized and relevant concepts of complex network theory and set it into the context of economic trade networks. After describing the mathematical formulation of complex networks (section 2.1), we apply the framework of networks to economic trade (section 2.2) and briefly discuss the main economical concepts of *input-ouput analysis* in section 2.3. Input-output analysis constitutes the conceptual basis of one of the databases that we utilize for the construction of trade networks in this thesis. In order to analyze the topology of trade, we require the application of specific methods and measures in complex networks. Next to commonly used and well-known network measures that are elaborated in more detail in comprehensive reviews [7, 8], we develop some newly defined and extended network measures such as the *corrected Hamming distance in weighted networks* in section 2.4. The notions of networks partitions, subnetworks and communities play an important role in understanding the substructure in economic trade networks and are described in section 2.5. Parts of the discussions in this chapter follow closely the presentations in P1 and P2.

2.1. Formulation of networks

A *graph* $\mathcal{G}$ is mathematically defined by a set of *vertices* $\mathcal{V}$ and a set of *edges* $\mathcal{E}$ [8]:

$$\mathcal{G} := (\mathcal{V}, \mathcal{E}) \, . \tag{2.1}$$

An edge $(i, j) \in \mathcal{E}$ can be interpreted as a connection between the two vertices $v_i, v_j \in \mathcal{V}$. In the context of physics, there is a more common usage of the term *networks* instead of graphs. Vertices are then referred to as *nodes* whereas edges are referred to as *links*. Throughout the rest of the thesis, we utilize this latter terminology. We denote the number of nodes by $|\,\mathcal{V}\,| = N$ and the number of links

by $|\mathcal{E}| = m$. A common representation of a network with nodes $\mathcal{V}$ and links $\mathcal{E}$ is the *adjacency matrix* $\mathbf{A}$ of dimension $N \times N$, defined as

$$(\mathbf{A})_{ij} := a_{ij} := \begin{cases} 1 & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{else} \end{cases} . \tag{2.2}$$

The network is called *undirected* if and only if the adjacency matrix is symmetric, otherwise we call it *directed*. For an easier distinction between the two types of networks, we denote symmetric adjacency matrices of undirected networks with an additional inverted comma, i.e. $\mathbf{A}'$ with $a'_{ij} = a'_{ji}\ \forall v_i, v_j \in \mathcal{V}$. In a directed network the i-th row of the adjacency matrix $\mathbf{A}$ lists all outgoing links of node v_i, while column j depicts all incoming links to node v_j. Links of nodes to itself, i.e. $a_{ii} = 1$, are referred to as *self-loops*. If the connection between nodes can be distinguished by different intensities, a specific weight is then attributed to each link in $\mathcal{E}$. Thus, the individual weights of the matrix $(\mathbf{W})_{ij} := w_{ij}$ describe the intensity of the link from node v_i to v_j in a *weighted* network. In *unweighted* networks all links have equal intensity and the consideration of $\mathbf{W}$ becomes redundant as in that case $\mathbf{W} \equiv \mathbf{A}$.

2.2. Trade networks

In trade networks the nodes represent economic entities such as business firms or, at a more aggregated and macroeconomic level, industrial sectors, whole countries or even regional trade blocks. Traditionally, the first approaches in applying tools from network theory to the analysis of global trade have defined the *International Trade Network (ITN)* as a network of countries that are linked to other countries by their import and export relationships [21–23, 31]. When a country v_i has exported to or imported from another country v_j, a link between the two nodes is drawn in the corresponding direction. If the monetary value of the exchanged goods is known, a weighted network can be constructed with link weights being proportionally defined to the exchanged trade volume.

With the compilation of multi-regional *input-output tables* [32, 33] (see also section 2.3) with a global coverage, newly available and reliable data allows for a more detailed view on the ITN by breaking down national economies into their sectoral composition. Multi-regional input-output (MRIO) tables distinguish the activities of the different industrial sectors in each country's economy and both the national and international trade relationships are depicted by the aggregated monetary value of the exchanged goods between two industrial sectors. Thus, a network can be constructed from an input-output table with one node representing one industrial sector of a specific country. As before, a directed link between two nodes is drawn with a weight w_{ij} proportional to the aggregated monetary value of the exchanged goods.

In contrast to the ITN on the sectoral and national level described above, there is no database that globally covers all trade relationships between individual business firms. However, in the case of Japan, private market research companies have collected data

that allow for the construction of a trade network on the business firm level [34]. Here, nodes represent individual companies and each company is further attributed to one specific industry. Due to business secrets and corporate interests the monetary value of a trade relationship are not available from this data. An unweighted link in this network therefore only depicts the existence and direction of a trade relationship between two firms.

2.3. Economic input-output analysis

As mentioned in the previous section, multi-regional input-output (MRIO) tables provide for a suitable database to construct the international trade network with a detailed resolution of industrial sectors [35]. The framework of input-output analysis has been first introduced by Wassily Leontief [36]. It is designed to capture the structure of the economy as a set of interdependent processes [37]. Due to its wide applicability for economic analyses and policy making [38], input-output accounting has been firmly established in national account systems to map the trade structure and flow of commodities within the national economies.

The output of the economy's producing entities, such as manufacturers and service providers, is either used for consumption by private households and governments or it is utilized as input by other industries for their respective operating businesses. The latter process, i.e. when some output of one industry is required as input for further production, is referred to as *intermediate trade* whereas the consumption of an end-user is described as *final demand*. The *value added* of industries includes among others the compensation of employees and the net operating surplus. Figure 2.1 illustrates the structure of an input-output table as standardized by the System of National Accounts of the United Nations [39]. The individual entries of the intermediate trade matrix $\mathbf{T}$ describe the monetary value of the goods that are exchanged between the different industries in the economy. The rows of $\mathbf{T}$ describe the supplier whereas the

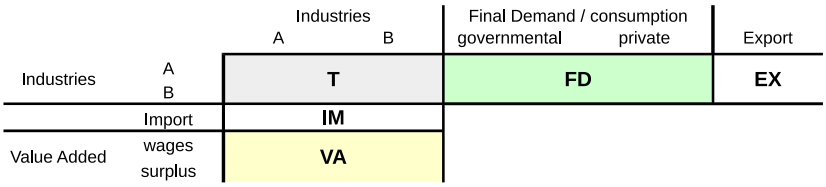


Figure 2.1.: Schematic illustration of the structure of a national input-output table with the block of intermediate trade $\mathbf{T}$, final demand $\mathbf{FD}$, value added $\mathbf{VA}$ and with imports $\mathbf{IM}$ and exports $\mathbf{EX}$.

columns describe the consumer of the produced good. Thus, the value of all goods that are sold from industry A (with row index 1) to industry B (with column index 2) is given by the matrix entry $(\mathbf{T})_{12}$ of Figure 2.1. In national input-output tables of one country, export and import blocks are attached next to the block of final demand $\mathbf{FD}$ and value added $\mathbf{VA}$, respectively.

The compilation of multi-regional input-output (MRIO) tables combines various data sources of national trade statistics, and international trade data from Eurostat, the Institute of Developing Economies, the OECD and the Comtrade database of the United Nations (for details, see [32]). MRIO tables have received rising attention in the last decade, with various analyses being conducted for consumer-responsibility accounting and for life cycle assessments of the economy [40]. In MRIO tables, the blocks for exports and imports become redundant, as shown in Figure 2.2. Here, the consumer of an industry's output that is distributed internationally is directly specified, distinguishing between the consuming industries and international end-users of the good.

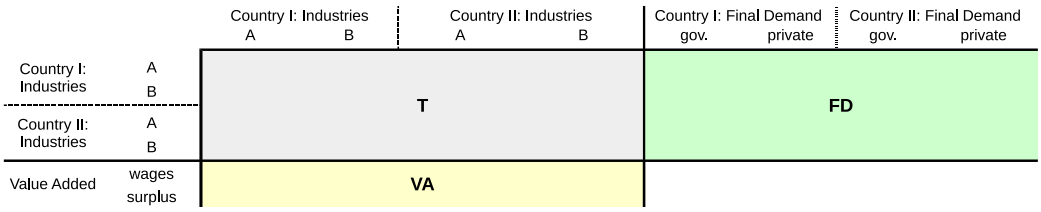


Figure 2.2.: Multi-regional input-output (MRIO) tables consist of a block of intermediate trade $\mathbf{T}$, final demand $\mathbf{FD}$ and value added $\mathbf{VA}$. In contrast to a national input-output table, international trade is included in the blocks of $\mathbf{T}$ and $\mathbf{FD}$.

The network interpretation of MRIO tables is straight forward. In order to construct a network $\mathcal{G}$, the set of nodes $\mathcal{V}$ is defined by the industries that are depicted by the rows and columns of the MRIO table. The structure of intermediate trade $\mathbf{T}$ then defines the set of links $\mathcal{E}$ and the weights $\mathbf{W}$ in the network. The final demand $\mathbf{FD}$ of the countries can be included in the set of nodes $\mathcal{V}$. By definition, these nodes have then only incoming flows of goods and services without any outgoing links. We further discuss the constructions of the investigated trade networks in detail in the corresponding sections of the respective chapter in this thesis.

2.4. Complex network measures

The theory of complex networks provides efficient methods and concepts to understand the topological properties of the system under study. One benefit of network theory is the applicability of its methods to a wide range of scientific disciplines. Here, we present the relevant network measures that are utilized throughout this thesis and discuss their interpretations in the context of trade networks. In general, network measures can be distinguished between their description of either *local* or *global* properties of the network. A network measure is referred to as local, if a value of this measure can be assigned to each individual node $v_i \in \mathcal{V}$ of the network. In contrast, a global network measure describes a topological property of the entire network.

2.4.1. Local network measures

Node degree and node strength In an undirected network with adjacency matrix $\mathbf{A}'$, the connectivity of a node $v_i \in \mathcal{V}$ is described by its *degree* $k_i = \sum_j a'_{ji} = \sum_j a'_{ij}$. In the directed case (with adjacency matrix $\mathbf{A}$), it is feasible to distinguish the in-degree k_i^{in} and the out-degree k_i^{out}, defined as [8]

$$k_i^{in} = \sum_{j=1}^{N} a_{ji} \qquad ; \qquad k_i^{out} = \sum_{j=1}^{N} a_{ij} \ . \tag{2.3}$$

The *total degree* $k_i^{tot} = k_i^{in} + k_i^{out}$ is then the sum of all incoming and outgoing links of a node. In weighted networks, the notion of the degree k_i is commonly extended to the *node strength* or *intensity* s_i. In undirected networks with weight matrix $\mathbf{W}'$ the strength reads $s_i = \sum_j w'_{ij}$. In the directed case (with weight matrix $\mathbf{W}$), one again distinguishes in-, out- and total strengths defined as [8]

$$s_i^{in} = \sum_{j=1}^{N} w_{ji} \qquad ; \qquad s_i^{out} = \sum_{j=1}^{N} w_{ij} \qquad ; \qquad s_i^{tot} = s_i^{in} + s_i^{out} \ . \tag{2.4}$$

In trade networks the direction of a link can depict either the direction of the monetary flow or the direction of the flow of goods. These two flows are always reverse to each other. In the ITN the link direction is often defined to correspond to the flow of goods. In this definition, the in-degree (out-degree) represents the number of suppliers (consumers) of one economic entity. If the data allows for the construction of a weighted network, the in-strength (out-strength) represents the value of all incoming (outgoing) goods traded by the economic entity.

Dependency measures If we consider the trade network as a flow network [41, 42] of goods and money, the *output dependency*

$$p_{ij}^{out} := \frac{w_{ij}}{\sum_{k=1}^{N} w_{ik}} = \frac{w_{ij}}{s_i^{out}} \tag{2.5}$$

describes the probability that a unit good in node v_i flows to v_j in one iteration of a flow model. This can also be considered as a random walk of a unit good in the network. On the other hand, in the process of obtaining inputs, the probability that a unit of money flows from node v_i to node v_j is given by

$$p_{ji}^{in} := \frac{w_{ji}}{\sum_{k=1}^{N} w_{ki}} = \frac{w_{ji}}{s_i^{in}} \ . \tag{2.6}$$

We utilize these measures and the picture of the flow network and random walk to define the measure of *trade interconnectedness* in chapter 4.

Equations (2.5) and (2.6) can also be interpreted in terms of a dependency relation between the economic entities v_i and v_j. The *input dependency* p_{ji}^{in} quantifies the

share of inputs of entity v_i which is supplied by entity v_j. A high value of p_{ji}^{in} implies that a large fraction of the inputs of entity v_i originates from v_j. Under the assumption that substitutions are costly and difficult to achieve due to additional market search costs, this would intuitively imply that v_i is highly dependent on v_j. The definition and motivation of the output dependency p_{ij}^{out} follows the analogous mechanism down the supply chain.

The input dependency shows similarities to the *Supply Propagation Connectivity* [43] that also measures the dependencies of economic flows. The two measures differ in the assumptions about the possibilities of the substitution of inputs: Whereas the inputs in the Supply Propagation Connectivity are assumed not to be substitutable from other industry sectors, here we do not distinguish between the industry attributes of the nodes in equation (2.6). Thus, we do not model the required composition of input factors of individual economic entities as being fixed. Instead, we allow for changes in the technology of the industries and focus on the fraction of the monetary value of the inputs that one economic entity buys from another.

Local clustering coefficient The *local clustering coefficient* measures the probability of the existence of a link between two randomly selected neighbors of a node v_i. In an unweighted and undirected network this probability C_i for node v_i is expressed by [8]

$$C_i = \frac{\sum_{i \neq j} \sum_{h \notin \{i,j\}} a'_{ij} a'_{jh} a'_{hi}}{k_i(k_i - 1)} \,, \tag{2.7}$$

where the notation implies a summation over all N nodes. The extension of equation (2.7) to weighted and directed network is ambiguous and different definitions of local clustering coefficients exist [44]. In this work, we follow the classification scheme of Fagiolo [45] and distinguish between the different flow patterns among the node v_i and its two neighbors. Figure 2.3 illustrates the four different flow patterns in which node v_i takes distinctive roles: In the pattern for the clustering coefficient C_i^{cyc} (see Figure 2.3A) there exists a cyclical relationship between the three nodes. Such a pattern arises in trade networks, when node v_i is a manufacturer that requires inputs from industries which themselves buy products that were manufactured with the output of v_i. C_i^{mid} captures the fraction of connected subnetworks where v_i has one incoming and one outgoing link but the relationship among the three nodes is not cyclical (Figure 2.3B). Both patterns C_i^{mid} and C_i^{cyc} describe the node's v_i importance as a transmitter of the flow of goods and money between the nodes. In contrast, the patterns of C_i^{in} (Figure 2.3C) and C_i^{out} (Figure 2.3D) indicate a node's role as sink and source, respectively. The mathematical formulation of these local clustering coefficients, that describe the fraction of closed links between the two neighbors, are

given by [45]:

$$C_i^{cyc} = \frac{(\tilde{\mathbf{A}}^3)_{ii}}{k_i^{in}k_i^{out} - k_i^{\leftrightarrow}} \quad ; \quad C_i^{mid} = \frac{(\tilde{\mathbf{A}}\tilde{\mathbf{A}}^T\tilde{\mathbf{A}})_{ii}}{k_i^{in}k_i^{out} - k_i^{\leftrightarrow}} \quad ; \quad C_i^{in} = \frac{(\tilde{\mathbf{A}}^T\tilde{\mathbf{A}}^2)_{ii}}{k_i^{in}(k_i^{in}-1)}$$
$$C_i^{out} = \frac{(\tilde{\mathbf{A}}^2\tilde{\mathbf{A}}^T)_{ii}}{k_i^{out}(k_i^{out}-1)} \quad ; \quad C_i^{all} = \frac{(\tilde{\mathbf{A}}+\tilde{\mathbf{A}}^T)_{ii}^3}{(k_i^{in}+k_i^{out})(k_i^{in}+k_i^{out}-1)-2k_i^{\leftrightarrow}} \,. \tag{2.8}$$

Self-loops are not considered in the clustering coefficient and equation (2.8) is valid, if all diagonal entries $\tilde{a}_{ii} = 0$ with $\tilde{\mathbf{A}} := \mathbf{A} - \mathrm{diag}(\mathbf{A})$. As introduced in equation (2.3), k_i^{in} represents the in-degree and k_i^{out} the out-degree of node v_i, while $k_i^{\leftrightarrow} = (\mathbf{A}^2)_{ii}$ denotes the number of bidirectional links associated with node v_i. The coefficient C_i^{all} does not distinguish between the different flow patterns and describes the fraction of neighbor pairs that are connected regardless of the direction of links. To adapt the concept of the clustering coefficient to weighted networks, we utilize the approach of [46] and replace the adjacency matrix $\tilde{\mathbf{A}}$ by $\hat{\mathbf{W}} = \mathbf{W}^{1/3} = \{w_{ij}^{1/3}\}$ with $\hat{w}_{ii} = 0$. If the weights $w_{ij} \in [0,1]$, then the clustering coefficients fulfills the condition $C_i^{\bullet} \in [0,1]$.

Motifs Closely related to the notion of the local clustering coefficient in directed networks is the concept of 3-node motifs. Motifs are defined as small induced subgraphs of the network and describe specific directionality patterns that are often attributed to fulfill specific functional roles in the network [47–49]. In general, one speaks of k-node motifs that consist of k nodes. In the case of 3-node motifs, 13 connected distinctive directionality patterns can be distinguished. We elaborate on motifs in more detail in chapter 6.

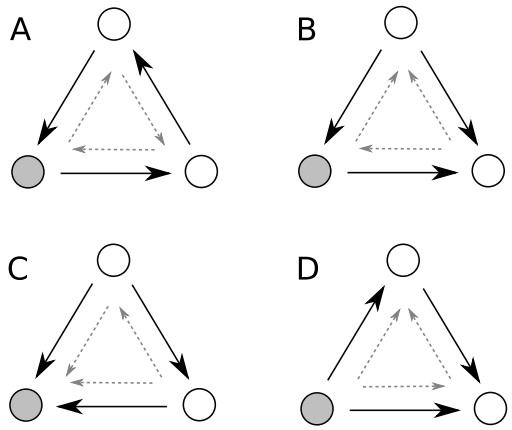


Figure 2.3.: Following the definitions in equation (2.8) the local clustering coefficients C_i^{cyc} (A), C_i^{mid} (B), C_i^{in} (C) and C_i^{out} (D) that refer to the gray, filled node v_i are shown. The two motifs that contribute to each coefficient are indicated by the three large black links and the three small gray links, respectively.

Betweenness The *betweenness* [50] measures the importance of a node with respect to its role as a mediator of the flow between nodes in the network. It is defined as [8]

$$b_i = \sum_{j \neq k \neq i} \frac{\sigma_{jk}(i)}{\sigma_{jk}} \,. \tag{2.9}$$

Here, σ_{jk} is the total number of shortest paths from node v_j to node v_k, while $\sigma_{jk}(i)$ is the number of these paths that include node v_i. In contrast to the previously discussed local network measures, the calculation of the betweenness of node v_i requires global information on the network.

In the considerations of this thesis, we neglect the weight information of links to calculate the shortest paths. Thus, we focus on the question whether trade relations between sectors have been established or not. A high value of the betweenness b_i implies that node v_i is important for the *global supply chain* in the economy. With the term global supply chain we describe the fact that individual supply chains cannot be obtained from the trade networks analyzed in this thesis. However, the links in the networks can be regarded as an aggregation of the individual supply chains. Hence, a node's position in the network can provide insights into the function of a node as an embedded entity in the global system of trade.

2.4.2. Global network measures

Link density The ratio between the existing number of links and the maximum number of possible links among the considered set of nodes is referred to as the *link density* [8]. Consequently, in the directed network the link density reads $\rho = m/N^2$, given that self-connections are considered in the network ($\rho = m/N(N-1)$ if self-connections are neglected). In trade networks, the link density describes the fraction of established trade relationships compared to a hypothetical fully-connected network in which every economic entity trades with every other economic entity.

Reciprocity In a directed network the *reciprocity* characterizes the probability that a randomly chosen link between two nodes also exists in the opposite direction. As self-connections do not provide additional information about this probability, flows of the node to itself are excluded [8]:

$$r = \frac{1}{m} \mathrm{Tr}\left[\mathbf{A} - \mathrm{diag}(\mathbf{A})\right]^2 \,. \tag{2.10}$$

Bidirectional links arise in trade networks, if v_i sells its output to v_j while the latter further produces goods that are required as inputs for v_i's production. For example, a renewable energy operator requires steel for the construction of its wind farms, whereas the heavy industry requires energy for the production of steel. This illustrative scenario describes a situation with a bidirectional link in a trade network.

Hamming distance The *Hamming distance* quantifies the dissimilarity between two networks $\mathcal{G}, \mathcal{G}^*$ that have the same set of nodes $\mathcal{V}$ but different sets of edges $\mathcal{E}$ and $\mathcal{E}^*$. It is originally designed for unweighted networks and defined by [29]

$$H(\mathcal{G}, \mathcal{G}^*) = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} |a_{ij} - a_{ij}^*|}{N^2} \ . \tag{2.11}$$

The principle is extendable to weighted networks. However, there are different possibilities to adjust the normalization factor of equation (2.11) in weighted networks. In this thesis, we introduce the following three novel variations of generalizations of the Hamming distance and compare their performance:

$$H_s(\mathcal{G}, \mathcal{G}^*) = \frac{1}{N^2} \sum_{ij} \frac{|w_{ij} - w_{ij}^*|}{w_{ij} + w_{ij}^*} \ ; \tag{2.12}$$

$$H_m(\mathcal{G}, \mathcal{G}^*) = \frac{1}{N^2} \sum_{ij} \frac{|w_{ij} - w_{ij}^*|}{\max(w_{ij}, w_{ij}^*)} \ ; \tag{2.13}$$

$$H_a(\mathcal{G}, \mathcal{G}^*) = \frac{1}{N^2} \sum_{ij} \frac{|w_{ij} - w_{ij}^*|}{\Phi} \ \text{ with } \Phi = \frac{\sum_{ij} w_{ij} + \sum_{ij} w_{ij}^*}{|\mathbf{A}| + |\mathbf{A}^*|} \tag{2.14}$$

and $|\mathbf{A}| = \sum_{ij} a_{ij}$. In H_a, the differences of link weights are normalized with respect to the average weight per link Φ in the two networks. We can further distinguish between different contributions to the Hamming distance of H_s and H_m according to [51]. This allows for a more detailed assessment of the dissimilarity between the two networks. Let b and c be the number of pairs that are linked in one network and unconnected in the other, with b counting the links in the network with higher link density ρ. Then, the Hamming distance H_m can be decomposed as follows:

$$H_m(\mathcal{G}, \mathcal{G}^*) = \Delta\rho + l_b + \Delta w_m := \frac{b-c}{N^2} + \frac{2c}{N^2} + \frac{1}{N^2} \sum_{ij} \frac{|w_{ij} - w_{ij}^*| a_{ij} a_{ij}^*}{\max(w_{ij}, w_{ij}^*)} \ . \tag{2.15}$$

Thus, three summands describing specific structural differences contribute to the Hamming distance: the link density difference $\Delta\rho = (b-c)/N^2$, the blinking links $l_b = (2c)/N^2$ [52, 53], and Δw_m summarizing the change in weights between pairs where both networks exhibit a link. When comparing two networks with different link densities, the existence of additional links that only exist in the network with the higher density is expected. Thus, the "corrected" Hamming distance accounts for this expectation. It is defined by neglecting the contributions that solely arise from the link density difference between the two networks: $H_m^* = H_m - \Delta\rho$. The definitions Δw_s and $H_s^* = H_s - \Delta\rho$ can be adapted analogously.

In order to illustrate the properties of the definitions in equations (2.14) and (2.15), we simulate modifications of a random graph [54] and measure the above proposed generalizations of the Hamming distance. We simulate an ensemble of random graph realizations in which the probability for the existence of a link accounts for 0.06 in

a network with 200 nodes. Furthermore, we distribute weights w to the links that are drawn from a log-uniform distribution with $w \in [1; 1000]$. Then, we modify the graph according to three possible rules:

1. select links randomly and shuffle their weights among the selected links

2. randomly select links and replace them to previously non-linked pairs of nodes

3. randomly create new links in the network

The links are randomly and uniformly selected in each modification. Figure 2.4 illustrates the results for modifications in which 10 % of the links have been relinked and 50 % of new links have been added to the network. We then vary the number of weights that are shuffled and denote the results for the various generalizations of the Hamming distance and their contributions. Each point is obtained from an ensemble of 50 realizations of the modifications.

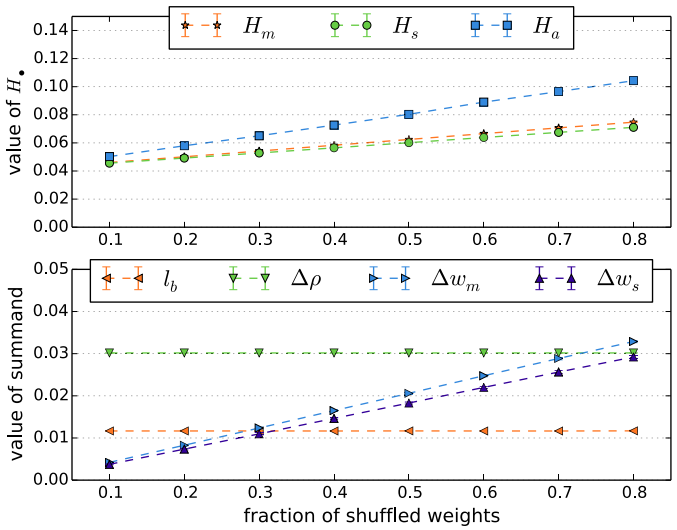


Figure 2.4.: Generalizations of the Hamming distance to weighted networks (top panel) in dependence of the amount of shuffled weights for modifications of a random graph with 10 % of replaced and 50 % of newly created links. The contributions of the difference in link density $\Delta\rho$, the blinking links l_b and the change in weights $\Delta w_\bullet$ is illustrated in the bottom panel.

We observe that contributions of the Hamming distance can be distinguished as described in equation (2.15). With an increasing probability to shuffle the weights, the contribution of Δw_m (Δw_s) to the respective Hamming distance H_m (H_s) increases.

Furthermore, we see that H_a increases to larger extent with a higher shuffling rate than H_s and H_m. This is due to the fact that H_a can be dominated by links with large weight differences, whereas this effect is balanced in the definitions of H_s and H_m. More specifically, each summand (i.e. pair of nodes) in equations (2.12) and (2.13) accounts for a value in the interval $[0, 1]$. A summand is 1 if a link from node v_i to v_j is present in $\mathcal{G}$ and absent in $\mathcal{G}^*$ (or vice versa). Therefore, H_s and H_m can be considered as an extension of the unweighted Hamming distance in equation (2.11) by additionally considering links that are present in both networks but have different weights.

Figure 2.4 is sensitive for the fraction of relinked pairs and newly created links. If the percentage of newly added links is increased, the contribution of $\Delta\rho$ increases and results in an upward shift of all the variations $H_\bullet$. On the other hand, if the fraction of relinked pairs is increased, the positive slopes of $H_\bullet$ and $\Delta w_\bullet$ become flatter whereas the level of l_b is shifted upwards and $\Delta\rho$ remains constant. The flatter slope can be explained by the fact that for networks that have only few links in common, the reshuffling of weights has only little effect as only few common links are available that can potentially contribute to $\Delta w_\bullet$.

The Hamming distance between two trade networks of different years quantifies to what extent the trade relations between the economic entities and the topological structure of the network have changed. In this context, a meaningful measure balances the domination of large weight differences as done in H_m and H_s. Therefore, they should be preferred over H_a. This is reasonable from an economic point of view: the establishment of new trade relations comes at higher costs and risks than the adjustment of the traded amount in existing trade relations.

2.5. Subnetworks and communities in networks

In this thesis, we analyze and discuss trade networks in which the individual nodes are business firms or industry sectors. On both scales the question arises, if specific groups of nodes are more interconnected among each other while the groups themselves are only sparsely connected to the rest of the network. The groups in such a structure are known as *communities* in network theory [55]. Thinking of international trade as a network of countries that all exhibit its own respective national economic network, this picture would coincide with the notion of a network out of interdependent subnetworks [56, 57]. However, in the process of an increasing globalized world economy this picture has been challenged in the literature [58].

To formally analyze the substructure of a trade network, we decompose the network into a *partition* $\mathcal{C}$ of subnetworks $\mathcal{G}'_p(\mathcal{V}'_p, \mathcal{E}'_{pp})$ that are induced by the subset of nodes $\mathcal{V}'_p \subset \mathcal{V}$ with $\cup_p \mathcal{V}'_p = \mathcal{V}$ and $\mathcal{V}'_p \cap \mathcal{V}'_q = \emptyset$ for $p \neq q$. Links can then be distinguished according to whether or not they connect nodes within the same subnetwork. Thus, the internal link sets $\mathcal{E}'_{pp}$ connect nodes that belong to the same subnetwork $\mathcal{G}'_p$,

whereas the cross-link sets $\mathcal{E}'_{pq}$ connect subnetworks via nodes that belong to the subnetworks $\mathcal{G}'_p$ and $\mathcal{G}'_q$, respectively. We further define the matrices

$$(\mathbf{A}_{\text{auto}})_{ij} = \begin{cases} 1, & \text{if } (i,j) \in \cup_p \mathcal{E}'_{pp} \\ 0, & \text{else} \end{cases} \quad ; \quad (\mathbf{A}_{\text{cross}})_{ij} = \begin{cases} 1, & \text{if } (i,j) \in \cup_{p \neq q} \mathcal{E}'_{pq} \\ 0, & \text{else} \end{cases} \tag{2.16}$$

that are convenient for the measurement of quantities that describe the internal subnetwork structure ($\mathbf{A}_{\text{auto}}$) or cross-subnetwork relations ($\mathbf{A}_{\text{cross}}$) with $\mathbf{A} = \mathbf{A}_{\text{auto}} + \mathbf{A}_{\text{cross}}$. A schematic illustration of a network with a subnetwork structure is depicted in Figure 2.5

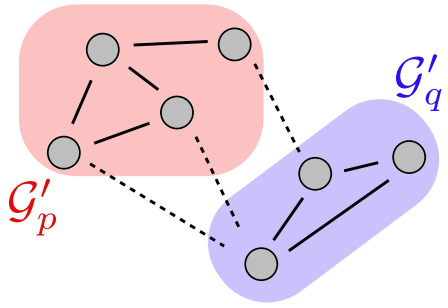


Figure 2.5.: Exemplary network illustration with two subnetworks $\mathcal{G}'_p$ and $\mathcal{G}'_q$. The solid lines represent the internal nodes of $\mathbf{A}_{\text{auto}}$ whereas the dashed lines are cross-subnetwork relations described by $\mathbf{A}_{\text{cross}}$.

A popular measure to assess the quality of a partition with respect to the notion of a community structure is the *modularity* [59]. The modularity Q is defined by the difference between the actual number of links within a community and the number that would be expected in a randomly linked network with the same degree sequence. For an undirected and unweighted network with adjacency matrix $\mathbf{A}'$ the modularity is defined as [59]

$$Q_u = \frac{1}{2m} \sum_{ij} \left(a'_{ij} - \frac{k_i k_j}{2m} \right) \delta(S_i, S_j) \,, \tag{2.17}$$

where k_i is the degree of node v_i, m is the number of links and S_i, S_j denote the indices of the communities that nodes v_i and v_j belong to. The Kronecker delta $\delta(S_i, S_j)$ assures that only node pairs within the same community contribute to the sum in equation (2.17).

Although various generalizations of the modularity exist, there is less consensus about the formulation of a generally applicable quality function for partitions in directed networks [60]. Arenas et al. [61] proposed to define the modularity in directed

networks as

$$Q_d = \frac{1}{m} \sum_{ij} \left(a_{ij} - \frac{k_i^{out} k_j^{in}}{m} \right) \delta(S_i, S_j) \ , \tag{2.18}$$

comparing the link distribution within a community to the expectation in the directed configuration model [62], with $k_i^{\bullet}$ as defined in equation (2.3). As suggested by Kim et al. [63] this approach does not fully account for the directionality of links between the nodes. Alternative definitions for the modularity are based on the attributes of links with respect to the probability density of the position of a random walker in the network [60]. However, in the context of trade networks, these definitions would lead to misleading interpretations that arise from the fact that industry sectors both create value added and request final demand. Thus, the monetary flow is not conserved in a trade network, which potentially traps a random walker in the sinks of the network. Therefore, we utilize the definitions of the modularity as defined in equations (2.17) and (2.18) in this thesis. By replacing the degree k_i^{in} (k_i^{out}) by the strength s_i^{in} (s_i^{out}) and the number of links m with the sum of weights $|\mathbf{W}| = \sum_{ij} w_{ij}$ in the network, the above definitions of Q are also applicable to weighted networks [61], i.e.

$$Q_{d,w} = \frac{1}{|\mathbf{W}|} \sum_{ij} \left(w_{ij} - \frac{s_i^{out} s_j^{in}}{|\mathbf{W}|} \right) \delta(S_i, S_j) \ . \tag{2.19}$$

In general there are multiple meaningful ways to partition a network. While community detection algorithms mainly seek to find a partition that maximizes a quality function (which is often the modularity), other partitions can be meaningful as well. In the ITN, for example, nodes can be grouped together according to their metadata, i.e. either according to their national or sectoral affiliation. In order to quantify the difference between two different partitions $\mathcal{C}$ and $\mathcal{C}'$, we measure the *variation of information* [64]

$$VI(\mathcal{C}, \mathcal{C}') = -\sum_{q=1}^{n_q} P(q) \log P(q) - \sum_{q'=1}^{n'_q} P(q') \log P(q') - 2 \sum_{q,q'} P(q, q') \log \frac{P(q, q')}{P(q) P(q')} \ . \tag{2.20}$$

Here, the probability that a randomly drawn node belongs to subnetwork $\mathcal{G}'_q$ in partition $\mathcal{C}$ with n_q subnetworks is denoted by $P(q)$. This probability can be expressed as $|\,\mathcal{V}'_q\,|/|\,\mathcal{V}\,|$. The joint probability that a random node belongs to subnetwork $\mathcal{G}'_q$ in partition $\mathcal{C}$ and to subnetwork $\mathcal{G}'_{q'}$ in partition $\mathcal{C}'$ is denoted by $P(q, q')$, i.e.

$$P(q, q') = \frac{|\,\mathcal{V}'_q \cap \mathcal{V}'_{q'}\,|}{\mathcal{V}} \ . \tag{2.21}$$

The value of VI is 0 if $\mathcal{C} = \mathcal{C}'$ and reaches its maximum value of $\log N$ in the case of $n_q = N$ and $n_{q'} = 1$.

We will see in the following chapter, how these introduced measures and concepts can be meaningfully applied to obtain new insights from a network of networks perspective on the ITN.

Chapter 3.

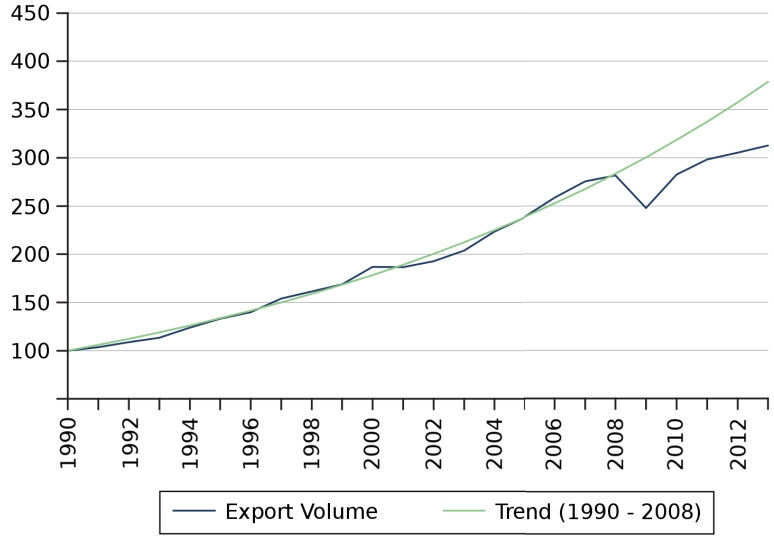
Trade on the global scale: World trade from a network of networks perspective

In this chapter, we focus on the specification of characteristic properties in the macroeconomic network of world trade, namely the International Trade Network (ITN). Mutually intertwined supply chains in the contemporary economy result in a complex interconnection and interdependence of trade relationships that entail a highly non-trivial network topology that varies with time. Here, we analyze the topological properties of the ITN on the scale of individual industry sectors. We construct the ITN for each year between 1990-2011 based on the information on intermediate trade in multi-regional input-output tables. In order to understand the complex interrelationships among different countries and economic sectors, as well as their dynamics, a holistic view on the underlying structural properties of the ITN is necessary. The aim of the analysis presented in this chapter is threefold: First, we provide a complementary view of the ITN to the literature by interpreting global trade as a network of subnetworks. A priori, there are multiple possibilities for the definition of the subnetworks. For example, the subnetworks can be defined by the national or by the sectoral affiliation of the nodes. We assess the meaningfulness of these two opposed views on empirical evidence and analyze the modular structure of the ITN in order to identify dominating communities in the world trade network. A second emphasis of this chapter consists of the identification of roles of specific industry sectors by focusing on the nodes' local properties in the network. Finally, we summarize the evolution of the ITN and discuss extreme events of trade reorganization within the observed time range during the process of increasing globalization. Parts of the following presentation in this chapter is closely related with the corresponding sections in publication P1.

3.1. Introduction

In the process of globalization, trade patterns have been reorganizing and international trade has been increasing almost continuously [65, 66]. Figure 3.1 shows the increase of internationally traded merchandise between 1990 and 2013. From the projection of the trend, we observe that the volume of international merchandise trade would have been likely to have tripled by 2009 if the financial crisis had not triggered the Great

Recession in this year. The crisis delayed the process by 2 years with merchandise trade having tripled by 2011 compared to its amount in 1990.



Source: WTO Secretariat

Figure 3.1.: Evolution of the volume of merchandise trade between 1990 and 2013. The volume of all merchandise exports in 1990 corresponds to the index of 100 [65].

To analyze the structure and reorganization of global trade relationships in this setting, the ITN has been traditionally defined upon the import/export relationships between countries. It has been analyzed as both a binary and a weighted, as well as a directed and an undirected complex network [22–24, 67]. Insights from these studies include that the network of world trade exhibits a distinctive non-random [21] and a core-periphery structure among countries [68]. More detailed analyses have focused on commodity-specific multi-network approaches [17, 69] that also address important aspects such as the community structure of the ITN [70, 71]. Previous research has proven that complex network theory provides a suitable toolbox in order to obtain a deeper knowledge about the economic principles of global trade. For example, from findings on local network centrality measures of individual countries in the ITN, the decline of the Western dominance in global trade is highlighted [72].

The literature on the ITN has mostly treated the countries as single nodes in the network. This approach neglects important substructures of the national economies. With the availability of multi-regional input-output (MRIO) tables [32, 35, 73] valuable and novel insights into the substructure of the ITN are obtained. Interpreted as a directed and weighted network, MRIO tables provide a more complete and highly resolved picture of the ITN in which the links depict the monetary value of the flow of goods and services between two industries. It allows for a more holistic view on the complex interdependencies within the present-day global economy. In particular, this refinement allows to address new questions on the topology of the network like

the following: How meaningful is the notion of national economies in an international globalized economy, where few transnational corporations hold dominant positions on a global scale [58]? What roles do specific industrial sectors and countries play in the ITN? In which industries and nations have trade relationships reorganized most along with globalization? How do national economies adapt to increasing foreign trade relations?

To find answers to these questions, we employ a MRIO database that comprises annually averaged monetary flows between 186 countries with 26 industrial sectors for the years 1990-2011. In this analysis, we focus on the interpretation of the ITN as a network of mutually interconnected subnetworks. As each node in this network is labeled with its country and industry, nodes can be intuitively grouped either by country or by industrial sector, defining a national and sectoral partition, respectively (see Figure 3.2). In addition, we aim to determine further data-driven partitions by utilizing established community detection algorithms [74]. The meaningfulness and the composition of the different partitions can then be quantified by the modularity and the variation of information.

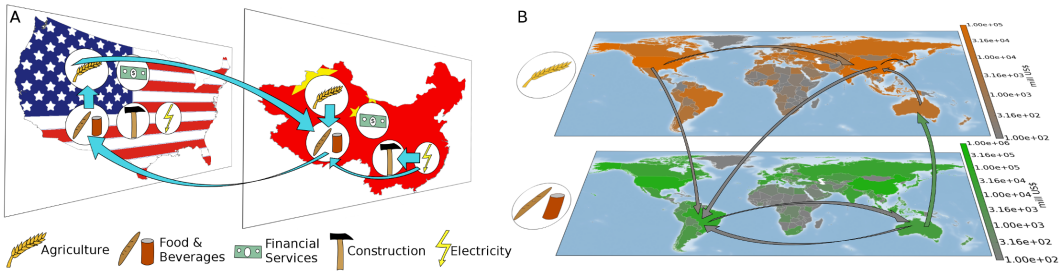


Figure 3.2.: Illustration of exemplary subnetworks in the ITN according to the national partition (A) and to the sectoral partition (B).

With the definition of partitions of the ITN, we utilize generalizations of the network measures that can be distinguished into internal and cross-subnetwork measures [56]. This approach allows for the identification of key players with respect to specific trade patterns and the assignment of roles to nodes in the ITN. In particular, we utilize extensions of the node strength, the clustering coefficient as well as the betweenness to the setting of interconnected subnetworks.

Finally, we address the implications of the globalization process on the structure of the ITN in the period between 1990-2011. We quantify trends in the respective network measures relating to internal or cross-linkages of subnetworks. We discuss how the introduced generalization of the Hamming distance to weighted networks serves as a suitable measure to quantify the inter-annual reorganization of trade patterns and illustrate its effectiveness in recognizing large-scale economic shocks and crises. Our results illustrate that the interpretation of the ITN as a network of networks exhibits new insights into the structural backbone of global trade and offers appropriate tools for the investigation of cross-sectoral economic relations at both the global and regional scale.

The remainder of this chapter is organized as follows: In section 3.2, we describe the utilized data and the construction of the ITN. The definitions of the subnetworks in the ITN as well as the generalizations of both local and global complex network measures to the network of networks framework is presented in section 3.3. The results of the analyses are presented in section 3.4. These include the modular structure (section 3.4.1), the attribution of roles (section 3.4.2) and the evolution of global measures in the ITN (section 3.4.3). We conclude the chapter with a discussion in section 3.5.

3.2. International Trade Network (ITN) construction

MRIO data summarizes the monetary value of the flow of goods and services between industrial sectors and can be meaningfully interpreted as a weighted and directed network of interdependent subnetworks, where nodes correspond to sectors (see also section 2.3). They allow for the construction of a network, in which the weighted and directed links describe the monetary value of the yearly aggregated flows of goods and money between the industries.

We utilize data from the Eora MRIO table providing annual data for 1990-2011 [32, 33] (Eora version 199.74). The Eora database collects highly resolved trade data of 186 countries, decomposing each economy of them into 26 industrial sectors. The industry classification is homogeneous in the sense that each country exhibits the identical classification. Thus, we construct for each year a network from the intermediate trade block $\mathbf{T}$ with $|\mathcal{V}| = N = 4836$ nodes. In this chapter we focus on the production supply chains and do not consider the final demand $\mathbf{FD}$ by the end-users. A list of all countries and of the homogeneous industry classification is given in appendix A. In Eora, the monetary value of the yearly aggregated flows of all goods and services between two industrial sectors are given in nominal US $. For the ITN in 1990, we consider two nodes to be linked if the monetary flow between two nodes exceeds 1 million US $. This is done to minimize effects from artifacts that arise from harmonization procedures during the compilation process of Eora [26, 27]. To minimize inflationary effects, we adapt the threshold to the yearly US inflation rate π_t [75] in the construction of the ITN for the following years. After the establishment of links, a weight proportional to the monetary value of the flow is then attributed to each edge. In order to distinguish structural changes from effects arising from inflation, we normalize the weights to the annual global trade volume for each year. Thus, we define the weight matrix $\mathbf{W}^t$ with the entries w_{ij}^t of the ITN in year t as

$$w_{ij}^t := \begin{cases} \frac{(\mathbf{T}^t)_{ij}}{\sum_{ij}(\mathbf{T}^t)_{ij}} & \text{if } (\mathbf{T}^t)_{ij} > 1 \cdot \pi_t \cdot 10^6 \text{ US \$} \\ 0 & \text{else} \end{cases} . \tag{3.1}$$

We further construct for each year a second network by fixing the amount of links to the number of links in the ITN in 1990. Thereby, we assess the robustness of the results with respect to varying the threshold during the network construction and

identify effects which are very sensitive to threshold variations in the construction process. The utilized threshold values of the ITN and the network with constant link density are shown in Figure 3.3. The trend in the threshold value for the network with constant link density implies a rise in trade volume in US $ and increasing entanglement in trade relationships.

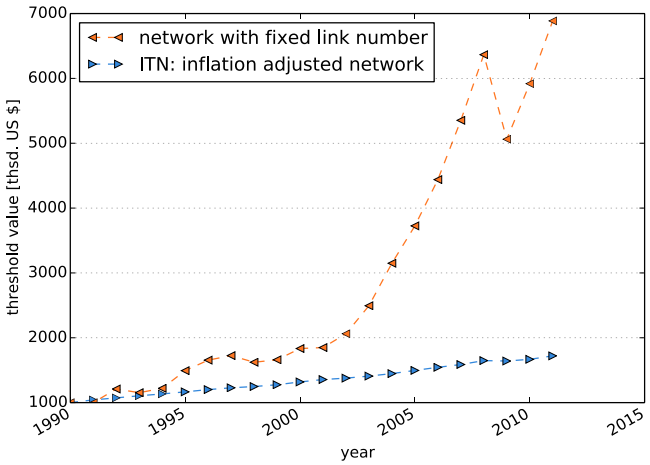


Figure 3.3.: The threshold values of the links weights for two different methods of network construction: The blue line illustrates the threshold value in the ITN that accounts for 1 Mio. US $ in 1990 with inflation taken into account. For the orange line, the network is constructed with the fixed number of links for all years.

3.3. Method generalizations in a network of networks

To empirically analyze the ITN with the interpretation of a network of subnetworks, we introduce partitions to the network and investigate both their internal and external topological characteristics. In the following, we present generalizations of the previously discussed concepts and measures of chapter 2 to set them into the context of interconnected subnetworks in the ITN.

3.3.1. Subnetworks in the ITN

In the ITN, the subsets of nodes $\mathcal{V}'_p \subset \mathcal{V}$ that induce the subnetworks $\mathcal{G}'_p$ can be meaningfully defined in various ways. Each node of the network belongs to a specific country c and to an industrial sector s. One of the self-evident partitions $\mathcal{C}_c$ is a classification of nodes by country. Figure 3.2A illustrates an excerpt of this partition, with the full network consisting of 186 national subnetworks containing 26 sectors each. This picture depicts the classical perception of economic trade theory in which

nations act as individual economic entities that organize their international trade relations with each other. A complementary approach is to define the partition $\mathcal{C}_s$ with subnetworks that consist of nodes from the same industrial sectors, as depicted in Figure 3.2B. The employment of dedicated community detection algorithms [74] provides a third reasonable way to partition the ITN. In general, most community detection algorithms are designed with the objective to find a partition with a maximized modularity score (equation (2.17)).

3.3.2. Local network measures on subnetworks

Node degree and node strength With a given partition $\mathcal{C}$, the degree of a node v_i (cf. equation (2.4)) can be further distinguished into the *internal degree* $k_{i;\text{auto}}$ considering all connections within the subnetwork of the node and the *cross-degree* $k_{i;\text{cross}}$. The latter describes the number of links to nodes of other subnetworks:

$$\begin{aligned} k^{in}_{i;\text{auto}} &= \sum_{j=1}^{N} \left(\mathbf{A}_{\text{auto}}\right)_{ji} \quad ; \quad k^{in}_{i;\text{cross}} = \sum_{j=1}^{N} \left(\mathbf{A}_{\text{cross}}\right)_{ji} \ ; \\ k^{out}_{i;\text{auto}} &= \sum_{j=1}^{N} \left(\mathbf{A}_{\text{auto}}\right)_{ij} \quad ; \quad k^{out}_{i;\text{cross}} = \sum_{j=1}^{N} \left(\mathbf{A}_{\text{cross}}\right)_{ij} \ . \end{aligned} \tag{3.2}$$

For a particular subnetwork with nodes $\mathcal{V}'_p$, we define the cross-in-degree of node v_i to this subnetwork as $k_i^{p;in} := \sum_{j\in\mathcal{V}'_p}(\mathbf{A}_{\text{cross}})_{ji}$ and the cross-out-degree as $k_i^{p;out} := \sum_{j\in\mathcal{V}'_p}(\mathbf{A}_{\text{cross}})_{ij}$. As in equation (3.2) for the degree, the strength can be analogously distinguished into the *internal strength* $s_{i;\text{auto}}$ and the *cross-strength* $s_{i;\text{cross}}$

$$\begin{aligned} s^{in}_{i;\text{auto}} &= \sum_{j=1}^{N} w_{ji} \left(\mathbf{A}_{\text{auto}}\right)_{ji} \quad ; \quad s^{in}_{i;\text{cross}} = \sum_{j=1}^{N} w_{ji} \left(\mathbf{A}_{\text{cross}}\right)_{ji} \ ; \\ s^{out}_{i;\text{auto}} &= \sum_{j=1}^{N} w_{ij} \left(\mathbf{A}_{\text{auto}}\right)_{ij} \quad ; \quad s^{out}_{i;\text{cross}} = \sum_{j=1}^{N} w_{ij} \left(\mathbf{A}_{\text{cross}}\right)_{ij} \ . \end{aligned} \tag{3.3}$$

The internal and cross-strength of the ITN as defined in equation (3.3) provide information about the trading partners of each node. For example in the national partition $\mathcal{C}_c$, the importance of a node for domestic and international trade is quantified by $s_{i;\text{auto}}$ and $s_{i;\text{cross}}$, respectively. Thus, in order to assess the characteristic trade patterns of a specific industry q, we aggregate the strength values of the industry over all countries for both the in-strength and out-strength, i.e. $s^{\bullet}_q = \sum_{i\in\mathcal{V}'_q} s^{\bullet}_i$. As we constructed the ITN with normalized weights (cf. section 3.2), $s^{\bullet}_q$ depicts the percentage of the total global annual trade volume that is traded by industry q in this context.

Local cross-clustering coefficient The concept of the *cross-clustering coefficients* that takes a given partition of a network into account has been introduced in [56]. In

the case of directed networks, we distinguish the cross-clustering coefficients further according to the different directionality patterns. Using

$$(\widetilde{\mathbf{W}}^{\mathbf{P}})_{ij} := \begin{cases} w_{ij}^{1/3} & \text{if } (i,j) \in \mathcal{E}'_{pp} \text{ and } i \neq j \\ 0 & \text{else} \end{cases} \tag{3.4}$$

$$\tag{3.5}$$

in networks with a defined subnetwork structure and k_i^p being the number of connections from node v_i to subnetwork $\mathcal{G}'_p$ (i.e. the cross-degree of v_i with respect to the nodes in $\mathcal{V}'_p$ [56]), the local cross-clustering coefficient for the "cycle" pattern yields:

$$C_i^{p;cyc} = \frac{(\hat{\mathbf{W}}\widetilde{\mathbf{W}}^{\mathbf{P}}\hat{\mathbf{W}})_{ii}}{k_i^{p;in}k_i^{p;out} - k_i^{p;\leftrightarrow}} , \tag{3.6}$$

with the bilateral cross-degree $k_i^{p;\leftrightarrow} = \sum_{j\in\mathcal{V}'_p} a_{ij}a_{ji}$. As in equation (2.8), $\hat{\mathbf{W}}$ describes the third root of the weight matrix with zeros on the diagonal. The other four distinct patterns of the directed clustering coefficient are adapted analogously:

$$C_i^{p;mid} = \frac{(\hat{\mathbf{W}}(\widetilde{\mathbf{W}}^{\mathbf{P}})^T\hat{\mathbf{W}})_{ii}}{k_i^{p;in}k_i^{p;out} - k_i^{p;\leftrightarrow}} \quad ; \quad C_i^{p;in} = \frac{(\hat{\mathbf{W}}^T\widetilde{\mathbf{W}}^{\mathbf{P}}\hat{\mathbf{W}})_{ii}}{k_i^{p;in}(k_i^{p;in}-1)} \quad ; \quad C_i^{p;out} = \frac{(\hat{\mathbf{W}}\widetilde{\mathbf{W}}^{\mathbf{P}}\hat{\mathbf{W}}^T)_{ii}}{k_i^{p;out}(k_i^{p;out}-1)}$$

$$C_i^{p;all} = \frac{[(\hat{\mathbf{W}}+\hat{\mathbf{W}}^T)(\widetilde{\mathbf{W}}^{\mathbf{P}}+(\widetilde{\mathbf{W}}^{\mathbf{P}})^T)(\hat{\mathbf{W}}+\hat{\mathbf{W}}^T)]_{ii}}{(k_i^{p;in}+k_i^{p;out})(k_i^{p;in}+k_i^{p;out}-1)-2k_i^{p;\leftrightarrow}} . \tag{3.7}$$

In the directed ITN the five directionality patterns of the clustering coefficients describe different roles of the node in the global supply chain (cf. section 2.4). We address the question to what extent industrial sectors show typical clustering patterns. Let $\mathcal{V}'_q$ be the subset of nodes that belong to industry sector q in $\mathcal{C}_s$ and $\mathcal{U}'_i$ be the subset of nodes in $\mathcal{C}_c$ belonging to the same country as node v_i. Then

$$C_q = \sum_{i\in\mathcal{V}'_q} \frac{C_i}{\sum_{j\in\mathcal{U}'_i} C_j} \tag{3.8}$$

is the sectoral mean value of the clustering coefficient C averaged over all countries. In order to avoid that the properties of the major economies dominate the results, equation (3.8) is normalized such that countries with high trade volume equally contribute to the average as countries with less trade.

Cross-betweenness By definition, the betweenness of a node provides an estimate of a sector's importance in the global trade network. A higher resolved picture is provided by the *cross-betweenness* [56] that is confined to geodesics between two

subnetworks. Generalizing the concept of betweenness to the context of a network of subnetworks, the cross-betweenness is defined as

$$b_i^{pq} = \sum_{j \in \mathcal{V}'_p, k \in \mathcal{V}'_q; k,j \neq i} \frac{\sigma_{jk}(i)}{\sigma_{jk}} \ . \tag{3.9}$$

Equation (3.9) quantifies the importance of node v_i to connect two subnetworks $\mathcal{G}'_p$ and $\mathcal{G}'_q$. Here, σ_{jk} is the total number of shortest paths from node v_j to node v_k, while $\sigma_{jk}(i)$ is the number of these paths that include node v_i. To calculate shortest paths, we neglect the weight information of links and focus on the question whether trade relations between sectors have been established or not. Note that shortest paths between nodes in $\mathcal{V}'_p$ and $\mathcal{V}'_q$ may pass through a third subnetwork. In particular, the values b_i^{pq} of nodes that do not belong to either of the subnetworks, i.e. $v_i \notin \mathcal{V}'_p \cup \mathcal{V}'_q$, contain vital information about the node's importance in connecting these subnetworks. Thus, to quantify how strong the two subnetworks $\mathcal{G}'_p$ and $\mathcal{G}'_q$ are interconnected among each other, we define the *cross-betweenness fraction* from nodes that belong to a third subnetwork:

$$B^{pq} = \sum_{i \notin \mathcal{V}'_p \cup \mathcal{V}'_q} b_i^{pq} \Big/ \sum_{j \in V} b_j^{pq} \ . \tag{3.10}$$

A low value of B^{pq} implies strong direct relations between the subnetworks $\mathcal{G}'_p$ and $\mathcal{G}'_q$ as most geodesics from nodes in $\mathcal{V}'_p$ to nodes in $\mathcal{V}'_q$ do not cross a third subnetwork. On the other hand, a value of $B^{pq} = 1$ indicates that the network does not contain direct connections between the subnetworks $\mathcal{G}'_p$ and $\mathcal{G}'_q$. To give an example, in the context of the national national partition in the ITN, this implies that two countries with $B^{pq} = 1$ do not trade directly with each other. On the other hand, a value of $B^{pq} = 0$ would imply a very strong interconnectedness between the two countries.

3.3.3. Global network measures with a subnetwork topology

Link density and reciprocity In a network of subnetworks further topological properties are revealed when we distinguish the global network measures by either taking only internal links or links that connect two different subnetworks into account. Then, the internal link density and the cross-link density are defined by

$$\rho_{\text{auto}} = \frac{\sum_p |\, \mathcal{E}'_{pp} \,|}{\sum_p |\, \mathcal{V}'_p \,|^2} \qquad ; \qquad \rho_{\text{cross}} = \frac{\sum_{p \neq q} |\, \mathcal{E}'_{pq} \,|}{\sum_{p \neq q} |\, \mathcal{V}'_p \,|\, |\, \mathcal{V}'_q \,|} \ . \tag{3.11}$$

Note that by definition of the modularity, the community detection algorithms seek to find a partition such that ρ_{auto} exceeds ρ_{cross}.

To analyze the contribution of the reciprocity with respect to the subnetworks, we distinguish between the inner and cross-reciprocity that are obtained by

$$r_{\text{auto}} = \frac{1}{|\mathbf{A}_{\text{auto}}|} \operatorname{Tr}\left[\mathbf{A}_{\text{auto}} - \operatorname{diag}(\mathbf{A}_{\text{auto}})\right]^2 \quad ; \quad r_{\text{cross}} = \frac{1}{|\mathbf{A}_{\text{cross}}|} \operatorname{Tr}\left[\mathbf{A}_{\text{cross}}\right]^2 . \tag{3.12}$$

Global cross-clustering coefficient In order to assess the structure of triangular linking patterns between subnetworks, a global perspective on the cross-clustering coefficient is required. With the local cross-clustering coefficient $C_i^{p;cyc}$ of node v_i as defined in equation (3.6), the associated global cross-clustering coefficient from subnetwork $\mathcal{G}'_p$ to subnetwork $\mathcal{G}'_q$ is defined as [56, 76]

$$C_q^{p;cyc} := \sum_{i \in \mathcal{V}'_q} C_i^{p;cyc} . \tag{3.13}$$

The same approach can be applied to all directionality patterns introduced in equation (3.7). Note that the relation in equation (3.13) is not symmetric, i.e. $C_q^p \neq C_p^q$. In the ITN, the global cluster cross-clustering coefficient provides further insights of how densely two industries or countries are connected with each other.

3.4. The network of networks topology of the ITN

3.4.1. Subnetworks & Communities

In the interconnected and global economy of today, the evident question arises of how meaningful the notion of a national economy still is [58]. We address this question by comparing the network topology of a national partition ($\mathcal{C}_c$) with both the topology of the complementary sectoral partition ($\mathcal{C}_s$) and the partition defined by a modularity maximizing community detection algorithm ($\mathcal{C}_m$). A priori all partitions have their own justification. On the one hand, domestic (internal) trade within a country is supported by a common policy framework, short geographical distances as well as the common cultural and patriotic identity of the people. Thus, policy implied trade barriers play in general a minor role in domestic trade. Furthermore, transportation and transaction costs between sectors in the same country are kept comparatively low. On the other hand, in the industry classification, many companies that are part of the supply chain of one product are aggregated to the same industrial sector. Therefore, we expect that for a multi-level production process of goods, complex supply chains result in high trading activity within the same sector. Thirdly, a partition that maximizes the modularity is well motivated by finding groups of nodes with a high internal connectivity and strong interconnectedness within one group.

We assess how the definition of the national partition $\mathcal{C}_c$ and the sectoral partition $\mathcal{C}_s$ coincides with the notion of communities in network theory [55]. For the community detection we start with considering the undirected and unweighted definition first

(equation (2.17)). We utilize a distinguished community detection algorithm to define the partition $\mathcal{C}_m$ and compare its quality with $\mathcal{C}_c$ and $\mathcal{C}_s$. Specifically, we employ the "multilevel algorithm" developed by Blondel et al. that extracts communities by a heuristic method based on modularity optimization [74]. This specific algorithm was tested to return a relatively high modularity at fast calculation times compared to other algorithms.

Two examples of communities of the partition $\mathcal{C}_m$ in the ITN of the year 2005 are listed in Table 3.1. We find that the "multilevel algorithm" preferably assigns nodes of the same country to the same community. Furthermore, strong economic interdependence resulting from geographical proximity or historical and political connections are represented in the community structure, e.g. most industries of France and Algeria are assigned to the same community. This example illustrates that the communities found by the multilevel algorithm tend to follow the national partition rather than the sectoral one.

Table 3.1.: Compositions of two selected communities in $\mathcal{C}_m$ in the ITN of 2005.

Community A		**Community B**	
# nodes	**country**	**# nodes**	**country**
26	Austria	25	Algeria
25	Hungary	24	France
24	Switzerland	1	Germany
24	Slovakia	1	Belgium
24	Slovenia	1	Luxembourg
23	Germany	1	Mauritania
20	Czech Republic	1	Czech Republic
5	Denmark	-	-
3	Russia	-	-
2	Netherlands	-	-
2	Belgium	-	-
2	Poland	-	-
2	UK	-	-
2	Italy	-	-
2	Sweden	-	-
2	Lithuania	-	-
2	Finland	-	-
2	Norway	-	-
2	Turkey	-	-
+ 16 countries with 1 node each		-	-

To further quantify this finding, we measure the variation of information (VI) (equation (2.20)) for the ITN for the years 1990-2011 (see Figure 3.4A). For all years, the national partitions $\mathcal{C}_c$ show the highest similarity with the partition of highest modularity, $\mathcal{C}_m$. Thus, a comparison between these two partitions allows for an identification of the strongest international trade relationships forming the backbone of global trade. To assess the significance of the similarity between $\mathcal{C}_c$ and $\mathcal{C}_m$, we compare the values of VI with those computed for the partition $\mathcal{C}_{m'}$ which is obtained

from a typical representation of the configuration model [77], i.e. from a random graph that obeys the same degree sequence as the original ITN. As expected, the partition $\mathcal{C}_{m'}$ of the network in which links are drawn at random differs significantly from $\mathcal{C}_c$. However, this behavior is not observed for the sectoral partition. In fact, as $VI(\mathcal{C}_s, \mathcal{C}_m) > VI(\mathcal{C}_s, \mathcal{C}_{m'})$, we conclude that $\mathcal{C}_s$ does not exhibit a modular link structure that is expected for communities in the traditional network theoretical sense. Therefore, our results indicate that trade relationships are not primarily established within the same industrial sector.

Taking link directions and weights into account, the modularity $Q_{d,w}$ as defined in equation (2.19) is shown in Figure 3.4B for all partitions in the ITN for all years. Our previously described findings are further supported by the fact that the modularity $Q_{d,w}$ is low for $\mathcal{C}_s$, whereas the values of $Q_{d,w}(\mathcal{C}_c)$ are in the range of modularity values obtained with the community detection algorithm. Over the 1990s, we observe a decreasing trend of both $Q_{d,w}(\mathcal{C}_c)$ and $Q_{d,w}(\mathcal{C}_m)$, whereas the modularity of $\mathcal{C}_s$ is rising except for the period of the global financial crisis in 2009.

To assess the impact of the weights in the modularity calculation, Figure 3.4C shows the modularity Q_d by considering the degree and neglecting link weights in equation (2.18). Here, the results show a decreasing trend in Q_d for all partitions. This decrease indicates an increasing entanglement of trade patterns - possibly due to a rising complexity, as partitions in trade patterns become less significant. In the trade network that is constructed with a constant link density for all years, this trend for Q_d is considerably weaker (see Figure 3.4D). The qualitative differences between Figure 3.4B and 3.4C indicate that industries with large trade volumes contribute significantly to the value of $Q_{d,w}$. From the comparatively high values of $Q_{d,w}$ we can conclude that industries with large trade volume are grouped within tightly connected communities. The difference $Q_{d,w}(\mathcal{C}_m) - Q_{d,w}(\mathcal{C}_c)$ increases slightly in Figure 3.4B. However, this difference does not exhibit marked changes over time when link weights are neglected (cf. Figure 3.4C,D).

To summarize the results presented above, our findings demonstrate that nations are still valid partitions in the sense of communities in complex network science. High trading industries build particularly tightly connected communities. However, the modularity shows a decreasing trend for all partitions in the ITN when link weights are neglected. This trend can be explained by new established links with comparatively low trade volume that cause a rising complexity of relationships within the global trade network. Interpreted in economic terms, these findings represent the increasing complexity in global supply chains.

3.4.2. Role assignment of nodes and industries in the ITN

To identify a node's functional role within the setting of the ITN as a network of subnetworks, the comparison between the internal topology of the subnetwork and the cross-subnetwork relations is of particular relevance. Certain nodes in the ITN often play a characteristic role in global supply chains. For example, the industries of some countries are specialized on the export of specific goods or resources. Thus,

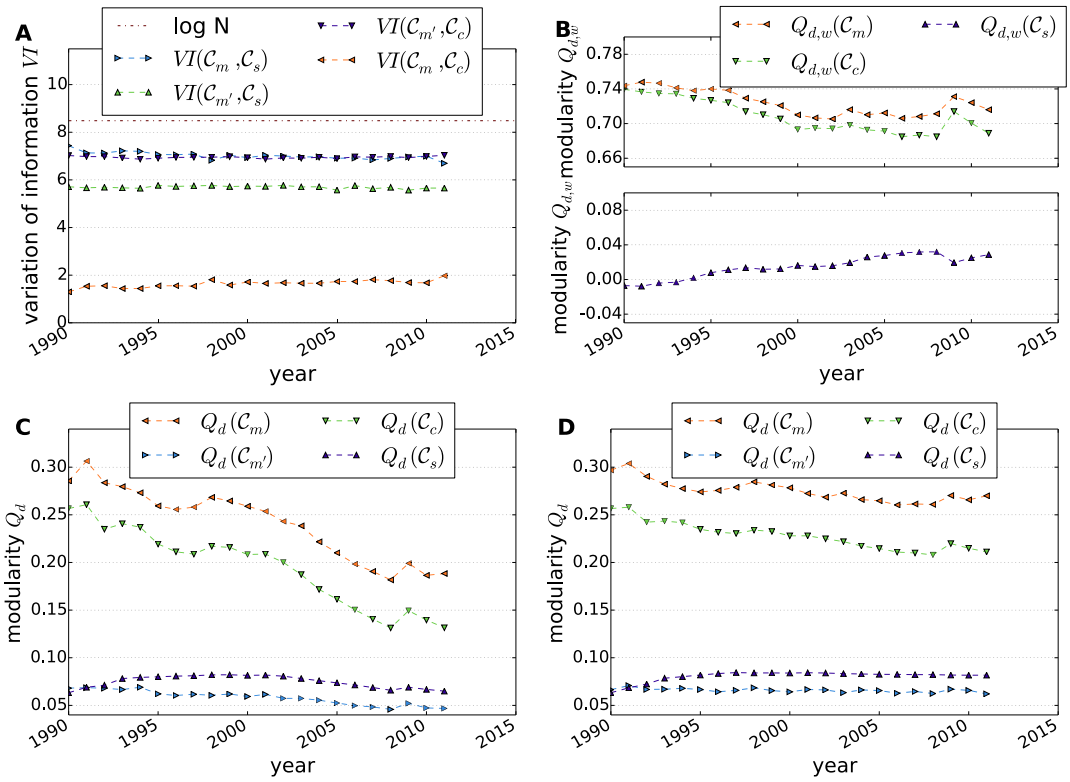


Figure 3.4.: Topological properties of different partitions of the ITN. (A) Evolution of the variation of information (VI) as similarity measure between two partitions of the ITN for 1990-2011. (B) Modularity $Q_{d,w}$ in the directed, weighted network for the partitions $\mathcal{C}_c, \mathcal{C}_s$ and $\mathcal{C}_m$ in the ITN. (C) Unweighted modularity Q_d for the partitions $\mathcal{C}_c, \mathcal{C}_s, \mathcal{C}_m$ in the ITN and for $\mathcal{C}_{m'}$ in the random graph with identical degree distributions. (D) Unweighted modularity Q_d for partitions in the ITN with constant link density.

from the interacting network perspective, these industry stand out as sources in the flow of goods across the subnetworks in the national partition. In order to identify key industries and to recognize their role in the global supply chain, we focus in the following on three network measures: node strength, (cross-)clustering coefficient and cross-betweenness.

Node strength The node strengths for sectors with the highest trade volume in the ITN of 2005 are summarized Table 3.2. We observe that financial services and business activities are particularly important for domestic trade. The domestic output of this industry amounts to 23.1 % of the total global trade. This corresponds to a share of 28 % of domestic trade while 81.9 % of all flows of goods and services are traded within national borders in 2005. The electrical and machinery industry holds the largest share of international trade, with $s^{in}_{q;\text{cross}} = 3.9$ %. Petroleum and chemical goods follow second in the ranking of cross-country trade. The complete tables that include the strengths of all industry sectors are provided in the appendix A.3 and A.4.

Table 3.2.: Key sectors for internal and cross-country trade in the ITN 2005. The values for the corresponding strength definitions are depicted in %. The corresponding exhaustive lists can be found in appendix A.3 and A.4.

	domestic input $s^{in}_{q;\textbf{auto}}$		**domestic output** $s^{out}_{q;\textbf{auto}}$
%	**industry**	%	**industry**
11.4	Financial Services & Businesses	23.1	Financial Services & Businesses
7.4	Electrical and Machinery	8.8	Petroleum, Chemical & Non-Metallic
7.1	Petroleum, Chemical & Non-Metallic	5.4	Transport
.	.	.	.
	foreign input $s^{in}_{q;\textbf{cross}}$		**foreign output** $s^{out}_{q;\textbf{cross}}$
%	**industry**	%	**industry**
3.9	Electrical and Machinery	2.9	Electrical and Machinery
3.2	Petroleum, Chemical & Non-Metallic	2.7	Petroleum, Chemical & Non-Metallic
1.6	Metal Products	1.4	Transport Equipment
.	.	.	.

The distributions of $s_{i;\text{auto}}$ and $s_{i;\text{cross}}$ for both the national partition $\mathcal{C}_c$ (A) and the sectoral partition $\mathcal{C}_s$ (B) are illustrated in Figure 3.5 for the year 2005. In $\mathcal{C}_c$ the distribution of the domestic flow of goods and services is shifted towards higher values compared to cross-country flows. This indicates that domestic trade is likely to exceed international trade for randomly drawn nodes. As there are more sectors abroad than in the same country of a node, this statement is even strengthened in significance when flows per potential trading partners are considered. Again, this supports the viewpoint of national economies being interconnected subnetworks in the ITN. In the sectoral partition $\mathcal{C}_s$ cross-sectoral trade exceeds intra-sectoral trade by absolute value. However, taking into account that there are more extra-sectoral nodes than intra-sectoral ones, trade within the same sector dominates cross-sectoral trade in terms of the monetary value of trade per potential trading partner. This

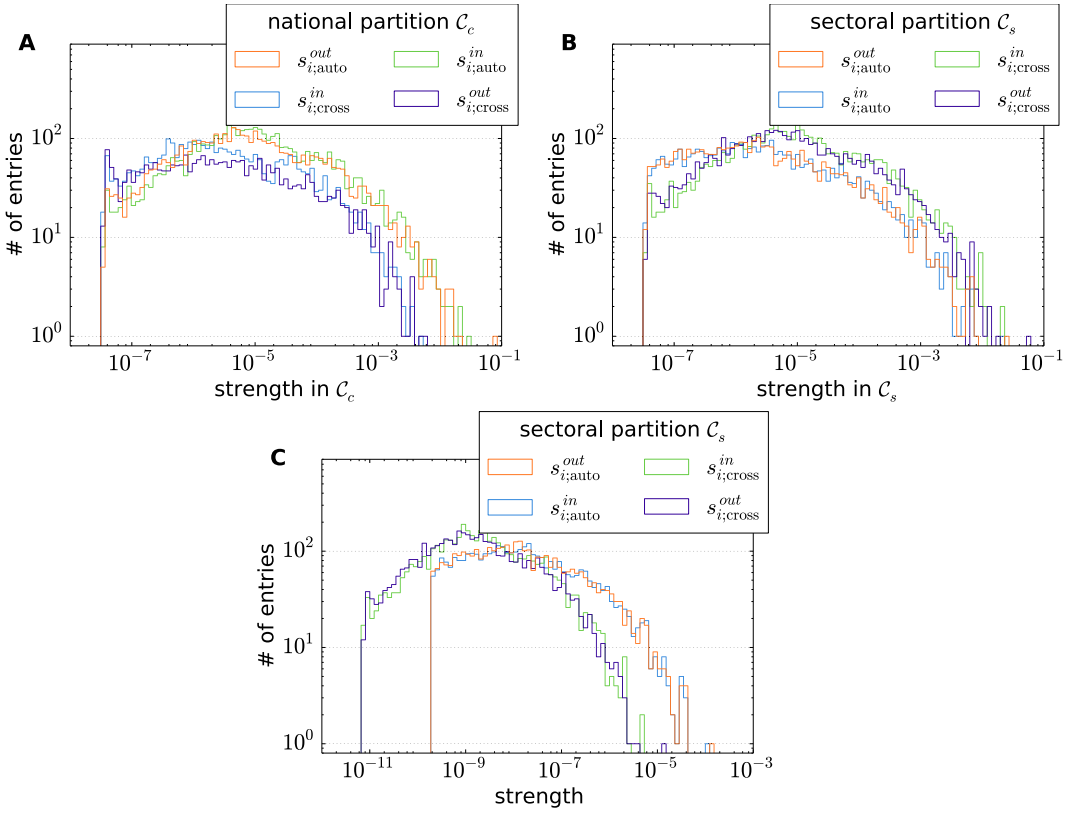


Figure 3.5.: Distributions of in- and out-strength in the ITN of 2005. (A) According to the national partition $\mathcal{C}_c$ and (B) to the sectoral partition $\mathcal{C}_s$. (C) Average strength per potential trading partner.

is illustrated in Figure 3.5C that depicts the distribution of the different strength definitions of equation (3.3) divided by the number of potential trading partners. With this normalization, we observe a shift of the distributions of $s^{\bullet}_{i;auto}$ to higher values compared to $s^{\bullet}_{i;cross}$.

Local and global cross-clustering coefficient We analyze the characteristic linking patterns of the industries in the ITN by investigating the cross-clustering coefficients of the industries' corresponding nodes. Figure 3.6 illustrates the results for the sectoral mean value of the local cross-clustering coefficients C_q as defined in equation (3.8) in the year 2005. We observe characteristic distributions for the different clustering coefficients. In particular, the motif C_q^{out} appears comparatively more frequent in nodes that belong to the sector of financial intermediation & business activities (cf. Figure 3.6D). This finding can hint towards the importance of the financial industry as capital provider for investments. Raw materials and resources are produced in the mining & quarrying industries and are often subsequently sold to other sectors, leading to a high rank in C_q^{out}. As shown in Figure 3.6B, the motif C_q^{mid} is frequently observed for sectors which are related to trade or which produce secondary products (e.g. petroleum, machinery). The construction industry is dominant in the motif C_q^{cyc} (cf. Figure 3.6A), whereas electrical and machinery industries dominate the pattern of C_q^{in} (cf. 3.6C).

The global cross-clustering coefficient as defined in equation (3.13) sheds light on characteristic trade patterns between subnetworks in the world trade network. We measure $C_q^{p;all}$ for all combinations of p and q in the national partition $\mathcal{C}_c$ and the sectoral partition $\mathcal{C}_s$. A summary of the highest obtained values is presented in Figs. 3.7A and 3.7B. In the national partition, the cross-clustering coefficient C_p^q is highest if $p = q$ for the world's largest economies. This is a reasonable behavior, as we have observed a high link density and trade volume in these national economies. Similar results are obtained for other directionality patterns of the clustering coefficient. As internal trade volume in subnetwork $\mathcal{G}'_p$ enters through a factor in the calculation of C_i^p (see equations (3.6) and (3.7)), subnetworks with a large trade volume exhibit large global cross-clustering coefficients. Therefore, the USA are involved in 20 of the 30 top global cross-clustering values in $\mathcal{C}_c$ (cf. Figure 3.7A). Furthermore, we conclude that the global cross-clustering coefficient is large between countries with high trade volumes and short geographical distance. For example, industries in Canada and Mexico score a high cross-clustering coefficient in the USA and the Netherlands, Belgium and France score high values in Germany (cf. Figure 3.7C). In $\mathcal{C}_s$ the electrical and machinery industry is the dominant sector (cf. Figure 3.7B).

Cross-betweenness To investigate the topological connectivity patterns between different subnetworks, we calculate the cross-betweenness fraction B^{pq} for each pair (p, q) of the 30 countries with the highest trade volume in the national partition $\mathcal{C}_c$ and for each pair of the 26 industry sectors in the sectoral partition $\mathcal{C}_s$. The distributions of B^{pq} in the national and sectoral partition are illustrated in Figure 3.8.

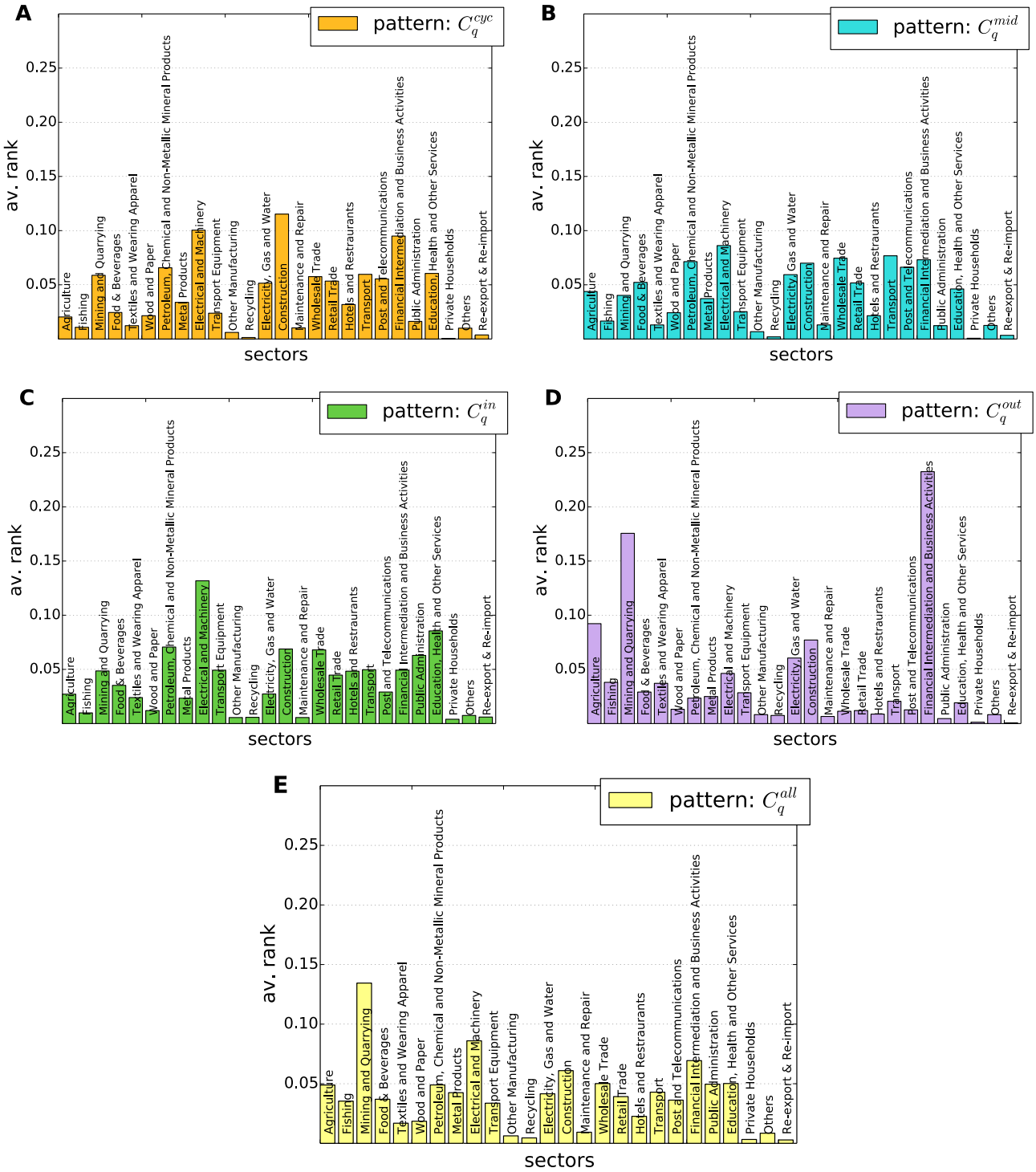


Figure 3.6.: Average rank of clustering coefficients C_q^{cyc} (A), C_q^{mid} (B), C_q^{in} (C), C_q^{out} (D) and C_q^{all} (E) as defined in equation (3.8) in the ITN of 2005.

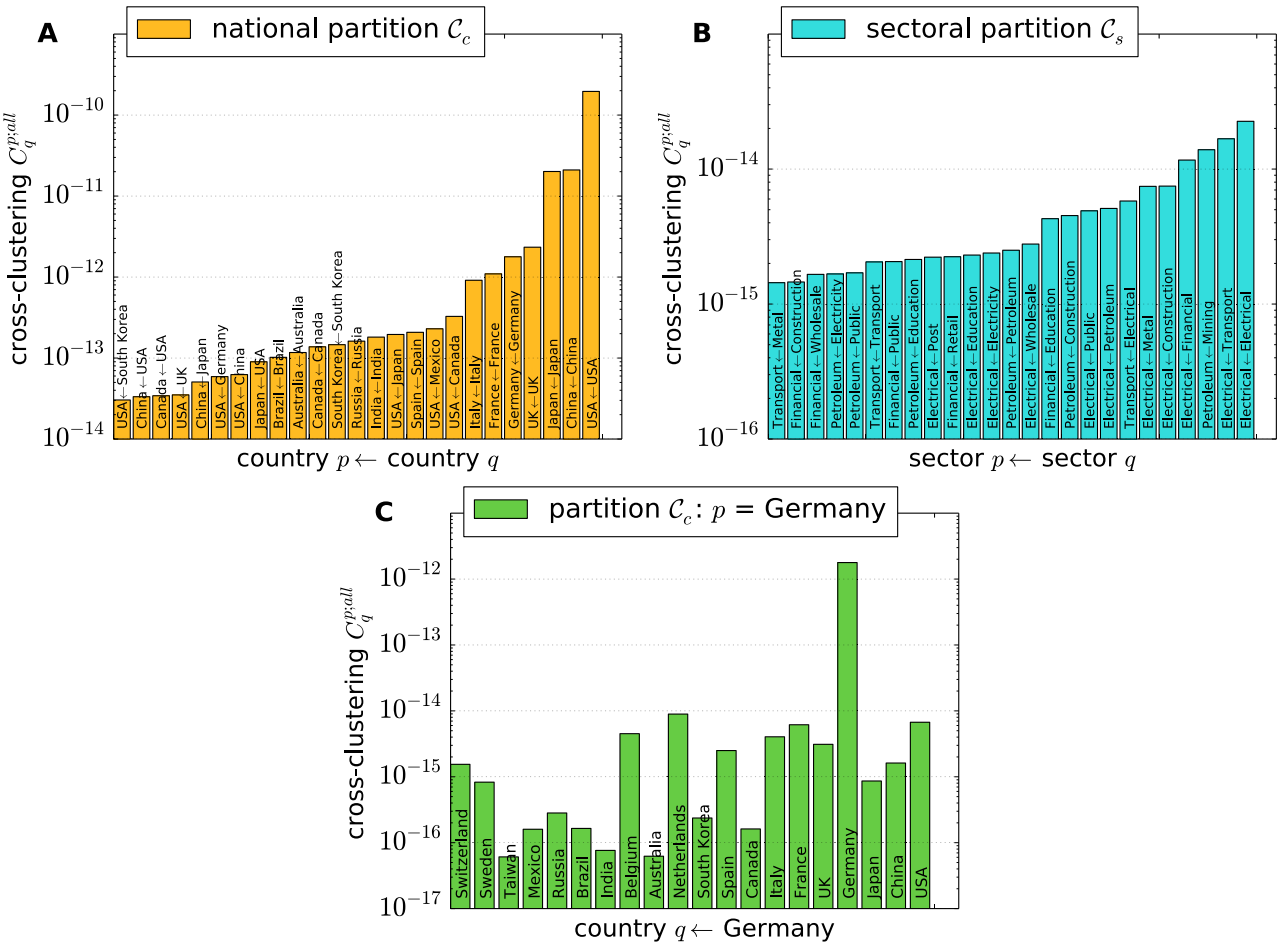


Figure 3.7.: Ranking of the largest global cross-clustering coefficients $C_q^{p;all}$ for pairs of countries (A) and sectors (B) in the ITN of 2005. (C) As an example, the global cross-clustering coefficients between Germany and the world's largest economies are shown.

We observe that the distribution of B^{pq} for the national partition is shifted towards lower values compared to $\mathcal{C}_s$. This is an indicator of a strong connectivity between national economies, as shortest paths between two countries often do not cross an additional third country. In fact, in the sectoral partition B^{pq} peaks at about 1. Thus, there are many shortest paths from sector p to sector q that run through at least one additional industry sector.

We are interested in identifying the countries and industries that play a significant role in connecting subnetworks $\mathcal{G}'_q$ and $\mathcal{G}'_p$. Thus, we calculate the average over pairs of subnetworks (p, q) taking only values b_i^{pq} into account if $v_i \notin \mathcal{V}'_p \cup \mathcal{V}'_q$:

$$\langle b_i \rangle = \sum_{p<q} \frac{b_i^{pq}(1-\delta_{ip})(1-\delta_{iq})}{\sum_{j \notin \mathcal{V}'_p \cup \mathcal{V}'_q} b_j^{pq}} \Bigg/ \binom{n_q}{2} \quad . \tag{3.14}$$

Here, δ_{ip} is defined as 1 if $v_i \in \mathcal{V}'_p$, and 0 otherwise. The number of subnetworks that are considered in the average is represented by n_q, normalizing equation (3.14) such that $\sum_{i=1}^{N} \langle b_i \rangle = 1$. The fraction in equation (3.14) describes the ratio between the cross-betweenness of node v_i and the sum of the cross-betweenness values over all nodes that belong to neither of the subnetworks p and q. The importance of subnetwork p in connecting other subnetworks is then obtained by $\sum_{i \in \mathcal{V}'_p} \langle b_i \rangle$. Our results show that for the national partition $\mathcal{C}_c$, Germany, USA and Switzerland hold most geodesics (Table 3.3). In the sectoral partition, however, USA, Germany and China lead the list with the financial services & business activities being the sector with largest $\langle b_i \rangle$ (Table 3.4 and Table A.6). We observe interesting differences between the geodesics in the national and sectoral partition. While some countries, in particular the larger economies like the USA and Germany, show a large cross-betweenness fraction in both partitions, some countries, e.g. Switzerland, belong only in one of the partitions to the most important subnetworks that connect two other subnetworks. The latter can occur, if a country focuses on international trade in one particular industry, i.e. on financial intermediation and business activities in the case of Switzerland.

Statistical interdependencies between local network measures A priori it is not known how the different measures introduced above contribute to complementary information about the network's topology and how the different measures are correlated with each other. In order to assess this issue, we investigate potential statistical interdependencies between the cross-strength, cross-clustering coefficient and cross-betweenness and present illustrative examples.

As link weights directly influence the calculation of the local cross-clustering coefficient in equation (2.8), we consider the correlation between $C_i^{p;all}$ and the monetary value of all exchanged goods and services $s_i^{p;out} := \sum_{j \in \mathcal{V}'_p} w_{ij}$ from node v_i into subnetwork $\mathcal{G}'_p$. We observe a stronger and generally positive statistical relationship between both characteristics in the national partition (cf. Figure 3.9A) than in the sectoral partition (cf. Figure 3.9B). The scatter plot between the local cross-clustering

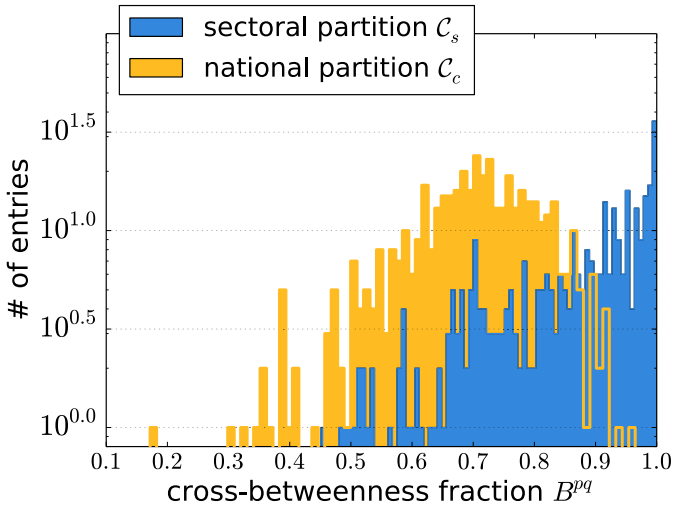


Figure 3.8.: Distribution of cross-betweenness fraction B^{pq} (cf. equation (3.10)) for pairs in the national partition $\mathcal{C}_c$ and the sectoral partition $\mathcal{C}_s$ of the ITN in the year 2005.

Table 3.3.: Largest values of the average cross-betweenness as defined in equation (3.14) in the ITN of 2005, aggregated by country. A longer list of countries is presented in the appendix A.5

National partition $\mathcal{C}_c$		**Sectoral partition $\mathcal{C}_s$**	
$\sum \langle b_i \rangle$	**country**	$\sum \langle b_i \rangle$	**country**
0.145	Germany	0.114	USA
0.103	USA	0.074	Germany
0.089	Switzerland	0.064	China
0.079	UK	0.038	France
0.066	China	0.036	Netherlands
0.061	Netherlands	0.032	Italy
0.050	Japan	0.030	Belgium
0.049	Italy	0.029	UK
0.043	France	0.025	Japan
0.041	Belgium	0.020	South Africa
0.145	Germany	0.066	China
0.103	USA	0.061	Netherlands
0.089	Switzerland	0.050	Japan
0.079	UK	.	.

Table 3.4.: Largest values of the average cross-betweenness as defined in equation (3.14) in the ITN of 2005, aggregated by industry. The complete table with all industry sectors is presented in appendix A.6

National partition $\mathcal{C}_c$		**Sectoral partition $\mathcal{C}_s$**	
$\sum \langle b_i \rangle$	**industries**	$\sum \langle b_i \rangle$	**industries**
0.178	Re-export & Re-import	0.193	Finance & Business
0.153	Petroleum	0.146	Petroleum
0.152	Finance & Business	0.115	Electrical and Machinery
0.119	Electrical and Machinery	0.107	Re-export & Re-import
0.060	Metal Products	0.098	Transport
0.047	Transport	0.060	Food & Beverages
0.043	Wood and Paper	0.045	Metal Products
0.038	Education	0.031	Education
0.035	Food & Beverages	0.027	Mining and Quarrying
0.028	Textiles	0.025	Agriculture
.	.	.	.

coefficient and the cross-betweenness exhibits a similar picture. In the national partition, shown in Figure 3.9C, the two measures show a stronger interdependence than in the sectoral partition in Figure 3.9D. This is due to the fact, that in the sectoral partition, one subnetwork consists of 186 nodes from countries with very different economic performances. This leads to a wide spread in the order of magnitudes in the weights w_{ij} and therefore to a wide spread of $C_i^{p;all}$, as the weights directly influence the clustering coefficient to the third power. However, the national partition exhibits fewer variability in the strengths of the 26 nodes that belong to the same country. We conclude that the cross-strength (measuring the overall monetary value of the flows originating from a node), the local cross-clustering coefficient (quantifying the occurrence of motifs across subnetworks), and the cross-betweenness (characterizing a node's importance in connecting two subnetworks) capture different aspects of a node's role in the ITN, although these three concepts are not fully unrelated conceptually.

3.4.3. Evolution of global interacting network measures in the ITN

As the third aspect of this chapter, we study the evolution of the ITN and how globalization is represented in the topological properties and substructures of global trade. The decreasing trend of the modularity in Figure 3.4B demonstrates that the community structure in the network has become less significant during the globalization process of the last decades. In the following, we discuss how the reorganization of trade patterns affects the network structure at both the local (node strength) and global scale (link density, reciprocity). We further investigate the speed of the reorganization process via the Hamming distance and discuss relevant measures to observe anomalies in trade patterns, in particular economic crises.

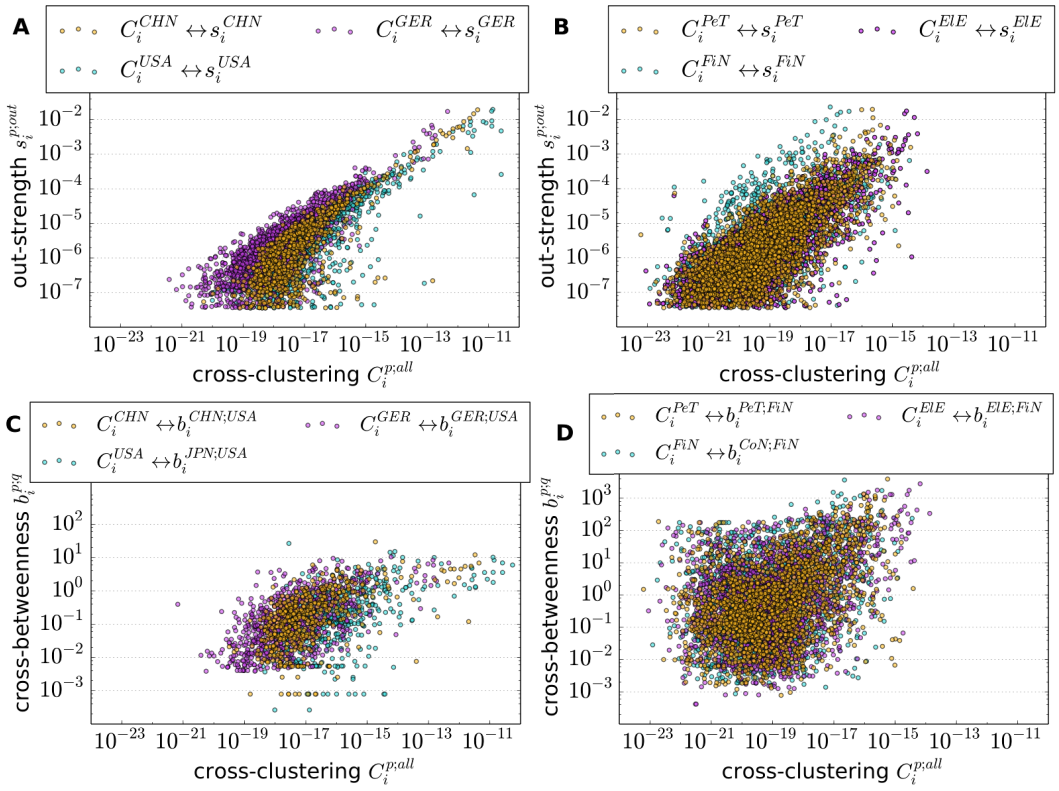


Figure 3.9.: Correlation between the network measures in different subnetworks (p, q). The scatter plots depict: (A) the local cross-clustering coefficient $C_i^{p;all}$ and the cross-strength $s_i^{p;out}$ for subnetworks in the national partition: China [CHN], Germany [GER] and USA. (B) $C_i^{p;all}$ and $s_i^{p;out}$ for the sectoral partition: Petroleum & Non-metallic products [PeT], Finance & Businesses [FiN], and Electrical & Machinery [ElE]. (C) $C_i^{p;all}$ and the cross-betweenness $b_i^{p;q}$ in the national partition: Japan [JPN], Germany, China and USA. (D) $C_i^{p;all}$ and $b_i^{p;q}$ in the sectoral partition: PeT, FiN, ElE and Construction [CoN].

Node strength To assess the evolution of the strength distributions presented in Figure 3.5 we calculate the mean of the node strengths for each year between 1990 and 2011. As we constructed the network by considering exclusively the intermediate trade block **T** and neglecting the final demand **FD** as sinks and value added **VA** as sources in the input-output tables, the means of the output and input distributions are identical. The opposing trends in the means of $s_{i;\text{auto}}$ and $s_{i;\text{cross}}$ are directly implied by the normalization step during network construction that we applied to avoid distorting effects arising from inflation and global developments. In Figure 3.10A, a trend towards more international trade can be observed from the evolution of the means in the national partition. However, in 2011 the mean domestic strength is still 4.4 times higher than the respective value for international relations. In the sectoral partition no comparable trend is observed with a practically stable mean. One could expect that technological progress leads to an adaptation of production functions towards the recent technologies, resulting in an adjustment of input requirements. However, in the classification of industry sectors these effects are small compared to the observed changes in the national partition.

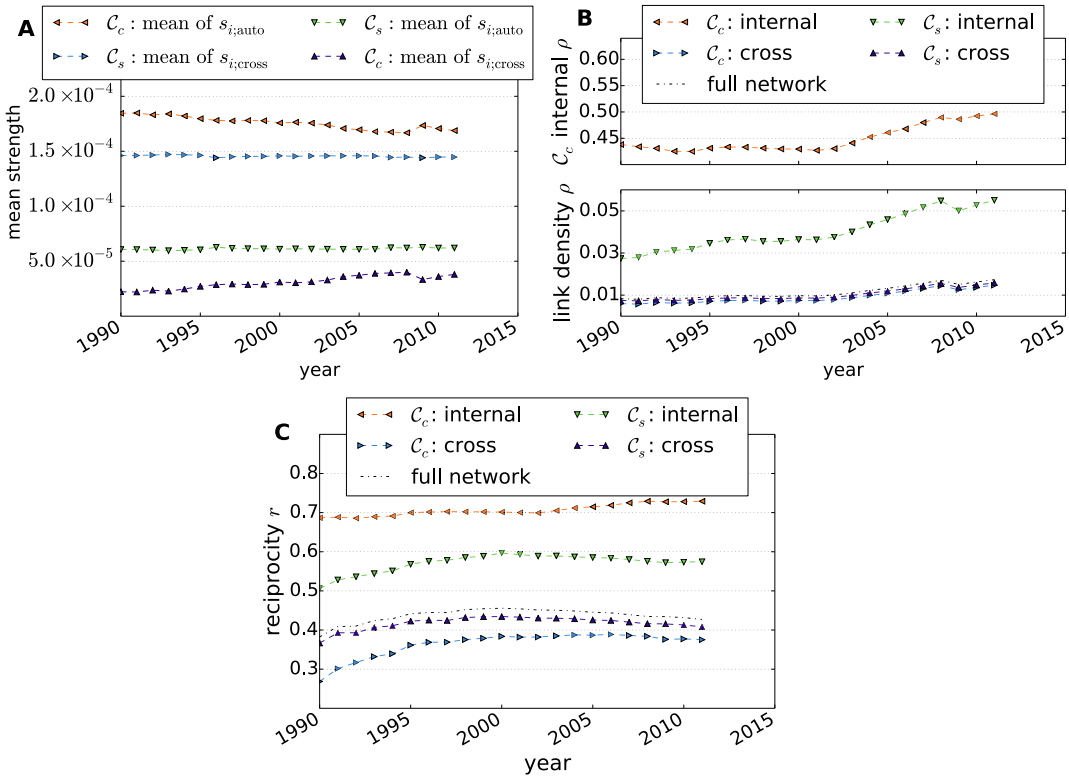


Figure 3.10.: (A) Mean node strength, (B) link density and (C) reciprocity. In addition to the properties of the full network, internal and cross-measures for the sectoral ($\mathcal{C}_s$) and national ($\mathcal{C}_c$) partition are shown.

Link density and reciprocity Regarding the impact of globalization on the global network measures, Figure 3.10B reveals an increasing trend in the link density in all parts of the network. The internal link density ρ_{auto} for domestic trade reaches 50 % in 2011, whereas the density for the full network ρ accounts for only 1.7 %. In the sectoral partition, ρ_{auto} exceeds ρ_{cross} with $\max(\rho_{\text{auto}}) = 5.5$ % in 2011. The results might be slightly biased due to the fact that data from national bureaus of statistics serve as main sources for the construction of the Eora MRIO database. This potentially leads to more accurate national data compared to the international flow of goods and services [32]. However, this bias is not able to explain the observed magnitude of differences in the link density between national and international trade. Therefore, our results further emphasize the importance of trade relations within national economies. A deviation from the trend of increasing link density is observed in 2009, when the link density decreased compared to the previous year. This effect coincides with the financial crisis in 2008/2009 that caused many countries to experience a recession in 2009 [78].

The reciprocity r (Figure 3.10C) exhibits a different behavior that depends on the considered partition. While r_{auto} gradually increases in the national partition $\mathcal{C}_c$, the reciprocity value peaks for the full network in the year 2000. This indicates that domestically, new links are mainly established between sectors that already possess a one-way trade relationship. However, reciprocity in cross-country relations saturates in 2000. For the full network and in the sectoral partition $\mathcal{C}_s$, reciprocity even decreases after 2000. This indicates, that in this period most emerging links are added as new one-way trade relationships between industrial sectors.

Hamming distance To quantify the restructuring of trade relations, we measure the Hamming distance between the ITN in the present and the preceding year. First, we compare the results of the different generalizations (equation (2.14)) of the Hamming distance (see Figure 3.11A). The graphs of H_m, H_s and H follow identical trends with H_m peaking in 2009 at a value of 0.0049. To better understand the underlying dynamics of the reorganization process, we measure the decomposition of the Hamming distance as defined in equation (2.15). We observe an increasing effect of link density differences since the year 2000 (see Figure 3.11B). In the corrected Hamming distance H_m^* the rising trend since 2000 is significantly reduced. Therefore, H_m^* is an applicable measure to identify anomalies in trade patterns, such as the financial crisis in 2009. Comparatively large values and fluctuations are visible in the early 1990s. These can partly be explained by an adaptation of trade patterns to the new global political and economic landscapes that arose after the collapse of the Soviet Union in 1991. Furthermore, the contribution of the difference in weights Δw_m to the Hamming distance increases compared to the blinking links l_b. For comparison, Figure 3.11C shows the Hamming distance of the trade network with constant link density for each year. In this network, the absolute values of H_m^* are lower than in the threshold based construction of the ITN. The peak in 2009 is still visible in the network with constant link density, although less significant compared to the

reorganization in the 1990s. We also measure the Hamming distances restricted to internal and cross-subnetwork connections for $\mathcal{C}_c$ and $\mathcal{C}_s$. However, differences in the trends are comparatively small, which implies that reorganization in trading patterns occurs in both internal and cross-subnetwork relations.

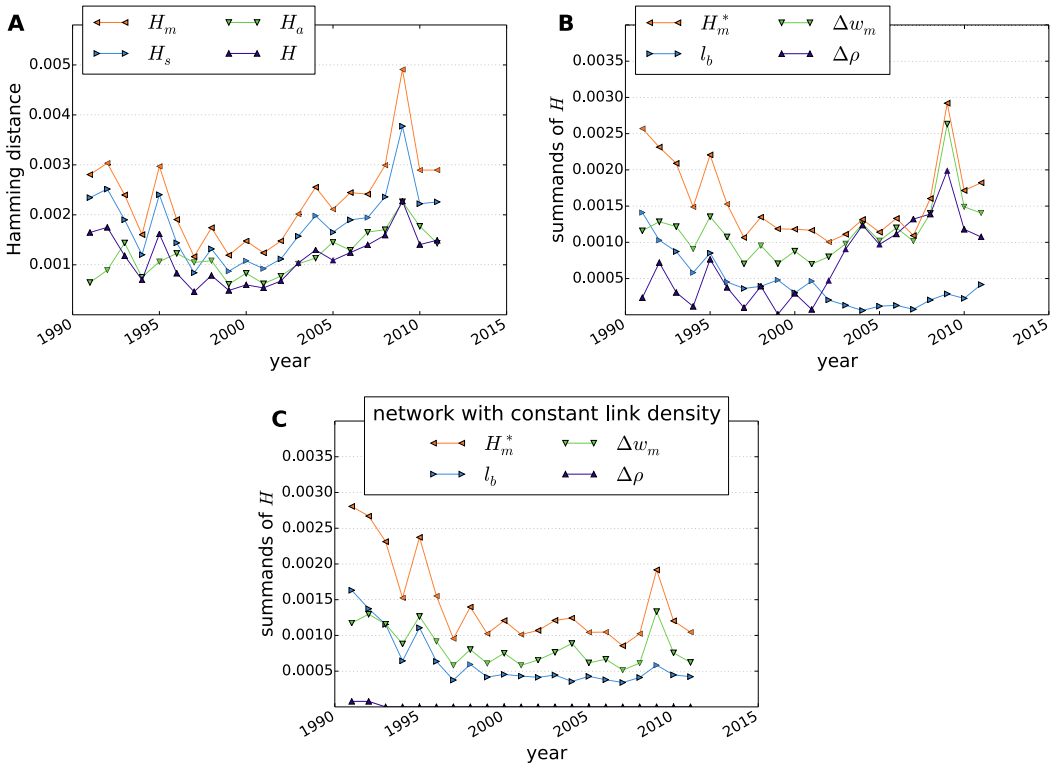


Figure 3.11.: Hamming distance between the ITNs of the current and the respective preceding year for 1991-2011. Various generalizations of H to weighted networks are compared (A). The contributions as defined in equation (2.15) are illustrated for the ITN constructed with thresholding (B) and constant link density (C).

3.5. Discussion

In this chapter we have empirically analyzed the topology of the ITN by interpreting the system of global trade as a network of interdependent subnetworks. With the definitions of different partitions of the network, the ITN reveals a non-trivial and dynamic architecture. The methodology and tools presented are well-suited for the assessment of both global and local properties of this network. In particular, we have shown how new insights can be obtained by the newly introduced measures of the cross-betweenness fraction and the generalizations of the Hamming distance to weighted networks.

In economic terms, we have addressed the question to what extent national economies and industries in the sectoral partition can be regarded as tightly connected economic regions. Our results demonstrate that the country-based partition of industrial sectors closely resembles the notion of communities in complex networks. However, an even higher modularity is achieved with a suitable community detection algorithm, pointing to an increasing relevance of international trade relationships. Important factors like geographical proximity and political linkages between countries are expressed in the observed community structure. A more detailed assessment of these factors and their implications for the network structure is an interesting subject for further study. Clusters in the sectoral partition do not exhibit the characteristic linkage features of communities.

The second key aspect of this chapter lies in the assignment of roles to the nodes in the ITN. Having defined meaningful partitions, the distinction between internal and cross-subnetwork properties provides a tool for unveiling the core functions of different sectors. For example, we find that domestic trade is dominated by financial services & business activities and that the trade activity of this sector accounts for > 20 % of global output. The clustering coefficient allows to assess directionality patterns and to find characteristic roles of sectors in the supply chain. Among others, the mining sector is identified as a predominantly output producing industry, whereas trade businesses appear more frequently in the center of global supply chains. Pairs of countries that are geographically close or exhibit large trade volume are often characterized by high cross-clustering coefficients. Further more detailed insights into the functional roles of industrial sectors and countries are provided by the cross-betweenness.

Finally, we have illustrated how globalization and economic crises have manifested themselves in the evolution of the substructure of the ITN. The increase of international interdependence is well observed in global network measures such as modularity and link density. The almost continuous decrease in unweighted modularity for both the optimal and the national partition suggests an overall increase in the complexity of trade relations, where the partition into national economies becomes less significant. However, the trends of most relevant network measures are interrupted in the year 2009, which coincides with the consequences of the global financial and economic crisis. We have successfully introduced a meaningful generalization of the Hamming distance to weighted networks that serves as a good indicator for the associated strong reorganization processes of trade patterns.

The methods and results provide a basis and reference for addressing more detailed research questions and case studies on trade networks. In the following chapter, we pursue the question to what extent bilateral trade agreements had an impact on the topological structure of the ITN. To this extent, we utilize methods that are motivated by the cross-betweenness of two subnetworks and introduce new appropriate network measures for this question.

Chapter 4.

The effects of bilateral trade agreements in the era of globalization

The number of negotiated bilateral trade agreements (BTAs) have increased considerably in the last decades. These BTAs influence the global trading system and the topology of the International Trade Network (ITN) as a whole. BTAs can potentially have far reaching impacts as the reorganization of flows of goods and services affects the contracting parties' industry sectors as well as other sectors that are directly or indirectly engaged in trade with these countries.

Here, we assess the impact of BTAs on the input-output linkages between the contractual parties' sectors by introducing the new measure of trade interconnectedness (TI). The TI captures the importance of how strong two subnetworks are interconnected among each other within the ITN. We then analyze the time evolution of the TI for each pair of countries with a BTA and define the BTA impact index that quantifies both the trend and changes in the level of the TI in the first years after the investigated trade agreements has come into force. After providing a global overview of the countries' BTA impact index, we focus in particular on the comparison of the trade profiles between two world leading economic powers, namely China and the USA. This chapter follows in parts closely the corresponding sections in publication P4.

4.1. Introduction

In the process of globalization, an increasing number of countries has been forming close trade relationships, often negotiated in reciprocal trade agreements. In these agreements, the countries grant each other trade privileges in form of concessions on trade barriers, which are for instance reductions of tariffs and quotas as well as easing of market access and of competition provisions. Theory suggests that the dismantling of trade barriers increases trade between the involved countries, which stimulates economic growth in both countries [79]. Empirical studies largely confirm a positive effect of BTAs on trade [80–82]. Yet, they also find that this positive effect might come at the cost of shifting production away from more efficient suppliers in other countries [82–84]. BTAs that are effective in enhancing bilateral trade might thus change the structure of the ITN formed by input-output linkages [25, 85–87]. Effectiveness is shown to depend on the countries' characteristics: for

instance being geographically close, sharing a common language and having a similar GDP is suggested to be beneficial in increasing trade gains [88–91]. Importantly, the effectiveness of BTAs is also determined by the scope and extent of commitments agreed on [92]. BTAs are very diverse in content and design, reflecting that they might be negotiated to serve other, strategic purposes as well [92, 93], e.g. underpinning politically motivated partnerships between countries, increasing bargaining power in trade negotiations with other countries and fostering liberal economic policy reform domestically [94]. Identifying and assessing the relative importance of the various possible objectives that might drive the negotiations proves difficult and are often a matter of interpretation of the final agreement [93]. In this context, two of the most active players in world trade are suggested to exhibit markedly different interests in forming bilateral trade agreement (BTA): While the USA is considered having a particular focus on rewarding their agreement partners for domestic economic reforms [95], China is assumed to be particularly interested in gaining economic and, indirectly, political influence by tying close economic dependencies [96].

Here, we investigate the effectiveness of BTAs by assessing their impact on the trade network spanned between the respective two countries. Gravity model analyses [97], which are often used for related analyses, focus in general exclusively on import and export positions. In contrast, we also take the topology of the ITN and higher orders into account as BTAs might also affect the demand and supply of sectors indirectly linked with the exporting and importing sectors. In particular, we define the measure of *trade interconnectedness (TI)* by investigating all direct and indirect input-output linkages within and across the respective two countries. Then, we assess the impact of BTAs by evaluating the time evolution of the TI, considering both the trend and changes in the magnitude of the TI. These methods are described in detail in section 4.2. In section 4.3, we analyze the effect of BTAs in general by drawing on 107 agreements that came into force between 1995 and 2008. Furthermore, we compare the results for the BTAs formed by the US and China and find empirical support for the suggested strategic differences in negotiating bilateral trade agreements. Section 4.4 concludes the chapter with a discussion of the results.

4.2. Trade interconnectedness and the BTA impact index

4.2.1. Trade interconnectedness

To analyze the interconnectedness between the economies of two countries in the ITN, we are motivated by the question how important the trade paths between two national subnetworks are for the respective countries. This notion is similar to that of the cross-betweenness as defined in equation (3.9) in section 3.3.2. However, here we are interested in a more comprehensive picture that exceeds the analysis of only shortest paths between the subnetworks. The consideration of longer paths is crucial to account for indirect effects and higher orders that are mediated by cross-sectoral supply chains. This represents the fact that two industries v_i and v_k that have no direct trade relation might nevertheless be dependent on each other by having an

identical trading partner v_j. In a scenario, in which v_j buys inputs from v_k and sells goods to v_i, v_i might be affected by a supply shortage of v_k which is mediated via v_j. Thus, to account for these higher order effects, we are here interested in the likeliness of a random walker to follow a specific path between the two countries in a normalized version of the ITN.

We construct the ITN from the multi-regional input-output database Eora as described in section 3.2. The utilized version of the data in this chapter (Eora version 199.82) covers the time range between 1990 and 2013. As policy makers are concerned about the final demand in their respective countries, we consider next to the intermediate trade block $\mathbf{T}$ of the input-output table also the final demand block $\mathbf{FD}$ in this chapter. The subset of nodes of one country consists then of the 26 industry sectors as shown in Table A.1 plus one node of the final demand. The final demand node is by definition an absorbing node of the flow of goods and services in the network that never exhibits any outgoing links.

From the ITN, we can deduce the dependency measures p_{ij}^{out} and p_{ij}^{in} as defined in equations (2.5) and (2.6) of section 2.4 to construct two flow networks. The weighted network with the entries p_{ij}^{out} can then be interpreted as the flow network of goods. In this network, the weight depicts the probability how likely a random walker follows the path from node v_i to node v_j. The probability corresponds to the ratio between the monetary value of the goods that have been sold from v_i to v_j and the monetary value of all of v_i's outputs within the investigated year. Similarly, in the flow network of money, the entries p_{ij}^{in} describe how likely a random walker follows the path of money from node v_j to v_i. Here, the probability is the ratio between the monetary value of the supplies that node v_j buys from v_i and all of v_j's inputs.

The matrix $\mathbf{P}_{out}$ with $(\mathbf{P}_{out})_{ij} := p_{ij}^{out}$ can therefore be interpreted as a flow network of goods, whereas $(\mathbf{P}_{in})_{ij} := p_{ij}^{in}$ describes a flow network of money. The powers of the matrix $(\mathbf{P}_{out}^{\lambda})_{ij}$ then describe the probability that the unit good follows a path of length λ from v_i to v_j. Similarly, $(\mathbf{P}_{in}^{\lambda})_{ij}$ is the probability that the monetary unit follows a path of length λ from v_j to v_i. We can utilize this concept to quantify the interconnectedness between the economies of two countries. This concept is based on the interpretation how likely it is that a random walker which starts from a node in one country finally arrives in the other country. Therefore, we utilize the national partition $\mathcal{C}_c$ of the network and group together sectors that belong to one country. The trade interconnectedness between the national subnetworks $\mathcal{G}'_x(\mathcal{V}'_x, \mathcal{E}'_{xx})$ and $\mathcal{G}'_y(\mathcal{V}'_y, \mathcal{E}'_{yy})$ of the two countries x and y is then defined as

$$TI^{\bullet}(\mathcal{G}'_x, \mathcal{G}'_y) = \frac{1}{|\,\mathcal{V}'_x\,| \cdot |\,\mathcal{V}'_y\,|} \sum_{\substack{v_i \in \mathcal{V}'_x \\ v_j \in \mathcal{V}'_y}} \left(\sum_{\lambda=1}^{\lambda_{max}} (\mathbf{P}_{\bullet}^{\lambda})_{ij} \right) , \tag{4.1}$$

with $\mathcal{V}'_x$ ($\mathcal{V}'_y$) denoting the subset of nodes that belong to country x (country y). The exponent λ describes the length of the path from node v_i to v_j in the sum of

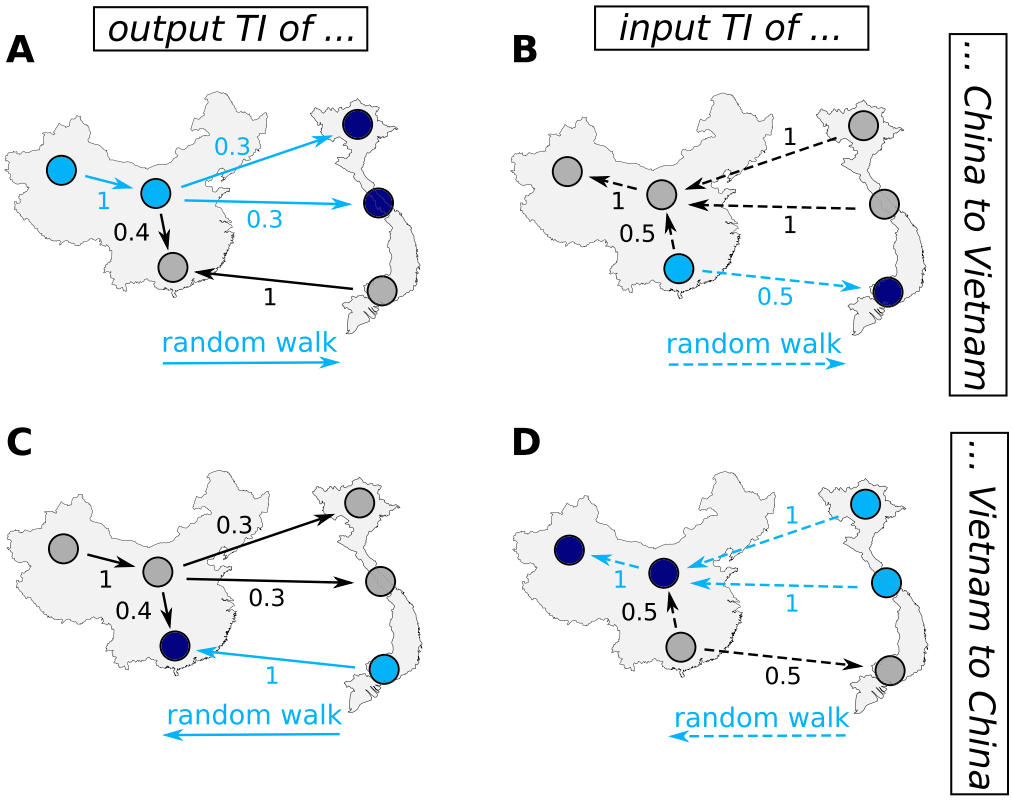


Figure 4.1.: Hypothetical excerpt of the ITN schematically illustrating the contributions to the different directions of TI. Colored circles indicate different industrial sectors, while solid (dashed) arrows indicate the flow of goods (money). In (A) the paths of goods that contribute to the output TI of China to Vietnam, TI^{out}(CHN, VNM), are highlighted in blue. Here, the unit good starts in China (light blue nodes) and ends in Vietnam (dark blue nodes). The individual path probabilities that are used to compute TI^{out} are the output dependency values p_{ij}^{out} which are illustrated by exemplary values at the links. In this example, paths of length one and two exist between the two countries (blue arrows). Supply directions that are not relevant for the supply of China to Vietnam are depicted by gray nodes and arrows. In (B), the paths of money that contribute to the input TI of China to Vietnam are marked blue, with the monetary flow following the opposite direction from the goods. The individual path probabilities used to compute the input TI are the input dependency values p_{ij}^{in}. The corresponding paths for the respective TIs of Vietnam to China are depicted in (C) and (D).

equation (4.1). We speak of $TI^{out}(\mathcal{G}'_x, \mathcal{G}'_y)$ as the output TI of $\mathcal{G}'_x$ to $\mathcal{G}'_y$ that can be interpreted as the relative importance of country y in the role of a consumer for country x. The relative importance of country x in the role of a supplier for country y can be analogously quantified with the input TI of $\mathcal{G}'_y$ to $\mathcal{G}'_x$, $TI^{in}(\mathcal{G}'_x, \mathcal{G}'_y)$. Note that $TI^{\bullet}$ is defined on a directed network and is not symmetric under the exchange of subnetworks, i.e. $TI^{\bullet}(\mathcal{G}'_x, \mathcal{G}'_y) \neq TI^{\bullet}(\mathcal{G}'_y, \mathcal{G}'_x)$. This is reasonable, as the importance of country y for country x is in general not the same as the vice versa relation. Figure 4.1 illustrates the paths that contribute to the different directions of TI on a hypothetical network between China (CHN) and Vietnam (VNM). The output TI of China to Vietnam, TI^{out}(CHN, VNM), accounts for the paths of goods that originate in China and end in Vietnam (see Fig. S4.1A). On the other hand, the input TI of China to Vietnam, TI^{in}(VNM, CHN), takes the paths of the monetary flow into account (see Fig. S4.1B). Here, the paths are defined in the opposite direction, as the money flows the opposite way from the good supply in the trade network. The corresponding paths of this exemplary network that contribute to the TIs of Vietnam to China are illustrated in Fig. S4.1C for TI^{out}(VNM, CHN) and Fig. S4.1D for TI^{in}(CHN, VNM), respectively. Notice that we do not consider paths that traverse a third country in the definition of $TI^{\bullet}$.

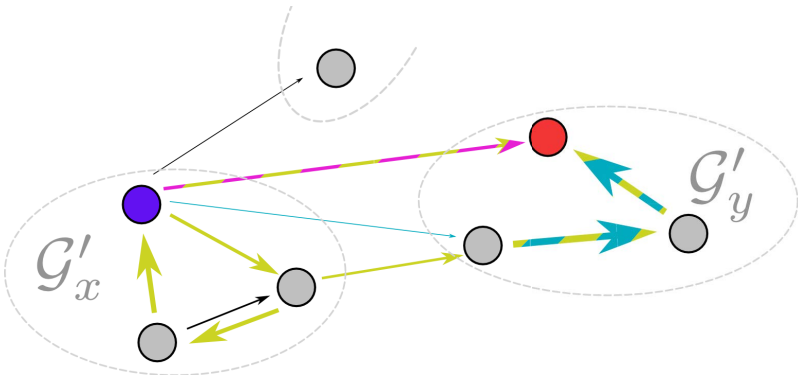


Figure 4.2.: Illustration of paths that contribute to $TI^{out}(\mathcal{G}'_x, \mathcal{G}'_y)$ between the countries $\mathcal{G}'_x$ and $\mathcal{G}'_y$ for $\lambda_{max} = 4$ for the exemplary pair of the colored nodes. The paths from the blue to the red node of length 1 (violet), 3 (light blue) and 4 (yellow) are highlighted. $TI^{out}(\mathcal{G}'_x, \mathcal{G}'_y)$ describes the sum of the random walk probabilities of these paths over all pairs of nodes with a start (blue node) in $\mathcal{G}'_x$ and an end (red node) in $\mathcal{G}'_y$. This is interpreted as the relative importance of $\mathcal{G}'_y$ as a consumer for $\mathcal{G}'_x$.

The maximal length of the paths that are to be considered in the TI is given by λ_{max} in equation (4.1). Figure 4.2 illustrates the paths between two nodes of an exemplary network excerpt that contribute to $TI^{out}(\mathcal{G}'_x, \mathcal{G}'_y)$ if $\lambda_{max} = 4$. By choosing $\lambda_{max} > 1$ in equation (4.1), higher orders that arise from cross-sectoral supply chains are considered in the measure. If λ_{max} is chosen to be dependent on the pair (v_i, v_j) and to represent the length of shortest paths between the pair of nodes, TI would consider the same paths as the cross-betweenness of equation (3.9). In the

case of $\lambda_{max} = \infty$, TI would resemble a version of the random walk betweenness as already discussed in the literature [98, 99]. In the context of this study, the contribution of high loop iterations that occur in $\lambda_{max} = \infty$ does not offer meaningful economic interpretations. Furthermore, in directed networks a random walker can be trapped within a certain subset of nodes resulting in divergences of the random walk betweenness.

The nodes of final demand take the role as sinks in the flow network of goods, causing a saturation in TI^{out} with increasing λ_{max}. However, these nodes become sources in the flow network of money and a converging behavior of TI^{in} is not observed. This is illustrated in Figures 4.3A and B that depict the distributions of TI^{out} and TI^{in}, respectively, for different choices of λ_{max} in the ITN of 2002. In these distributions all pairs of countries are considered that have negotiated a bilateral trade agreement in the investigated time range (see also section 4.3). It can be seen, that the values of TI^{in} do not convergence for economically reasonable path lengths.

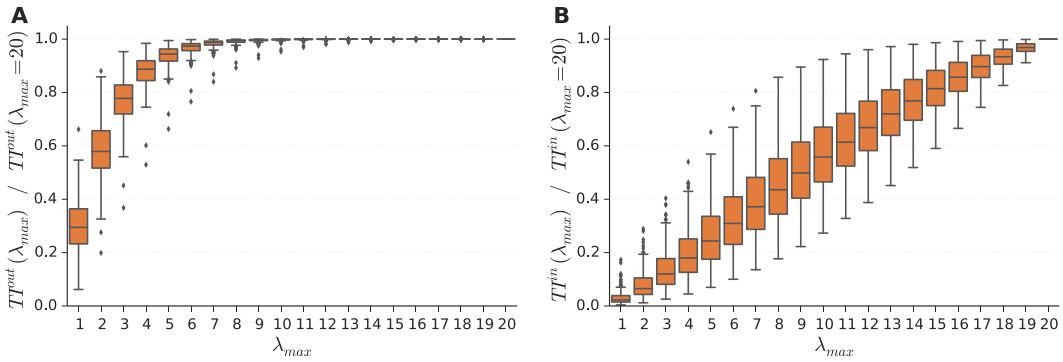


Figure 4.3.: Box plot of the distributions from all country pairs with a BTA of the TI^{out} (A) and TI^{in} (B) for different choices of the maximal considered path length λ_{max}. In both panels the ratio of $TI^{\bullet}$ to a reference maximal path length of $\lambda_{max} = 20$ are depicted. The box plot shows the median, upper and lower quartiles as well as outliers of the distribution. Outliers are defined if they exceed 1.5 times the inter-quartile range.

Thus, loops within one country of the trade network gain in importance for the TI^{in} with increasing maximal path lengths λ_{max}. The probabilities of these national loops decrease with time, as international trade has increased in the investigated time period (see section 3.4.3). As an example of a typical time series of TI^{in}, we investigate the time evolution of TI^{in} of Algeria to the European Union as displayed in Figure 4.4. We observe that with increasing maximal path length, the trend in the input TI decreases, as national loops are less probable in more recent years.

The above discussion illustrates that the maximal path length λ_{max} cannot be chosen arbitrarily large, since otherwise loops within one country would be overrepresented in TI^{in}. On the other hand, we have seen that higher-order paths affect the TI and should be considered in the analysis. In view of this trade-off, we decide to set λ_{max} to twice the average path length between the two subgraphs $\mathcal{G}'_x$ and $\mathcal{G}'_y$ of

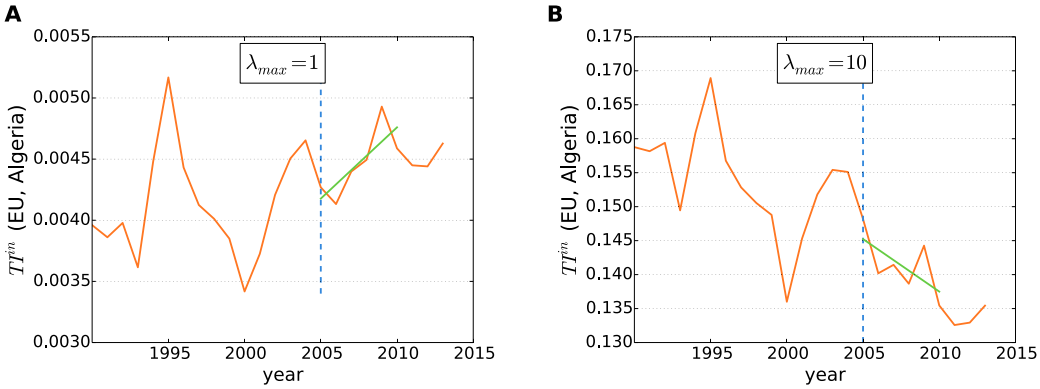


Figure 4.4.: The input TI of Algeria to the European Union for (A) $\lambda_{max} = 1$ and (B) $\lambda_{max} = 10$. The year in which the BTA comes into effect is indicated by the blue vertical line. The regression model selected by the AIC criterion and the corresponding maximum likelihood fit is displayed by the green line.

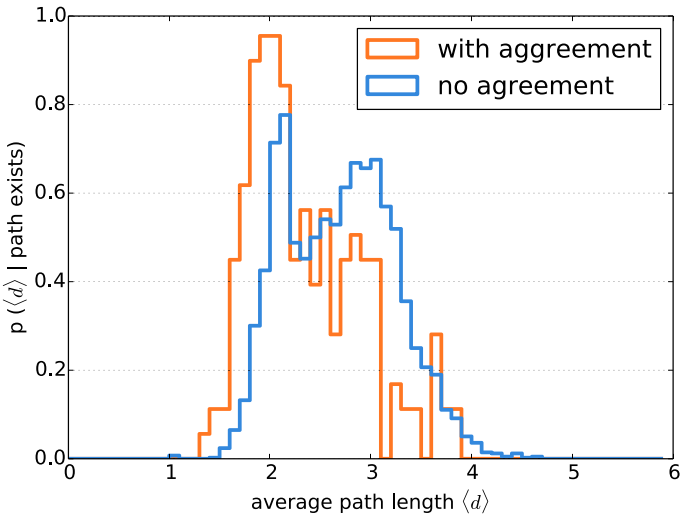


Figure 4.5.: Probability distributions of the average path length for pairs of countries with a bilateral trade agreement (orange) and without any agreement (blue) given that a direct path between the countries exists.

the ITN. This choice allows the consideration of sufficiently high-order paths between these subgraphs while at the same time avoiding too large contributions from loops within any of them. Figure 4.5 shows the probability distribution of the average path length $\langle d \rangle$, distinguished between the set of country pairs with and without a trade agreement and given that a direct path between the countries exists. We observe a clear shift to longer average path lengths for country pairs without agreement. The explicit values of the average path lengths between all pairs of countries with a BTA negotiated in the considered time period is given in the appendix (Table B.1). In appendix B, we also discuss the sensitivity of the following results in this chapter with respect to the choice of the maximal path length λ_{max}. We observe that the main conclusions of the results presented below are robust with respect to the choice of λ_{max} for economic reasonable path lengths (i.e. $\lambda_{max} \leq 10$).

4.2.2. BTA impact index

To quantify the impact of a BTA on the trade interconnectedness between the economies of the involved countries, we analyze the time series of *TI* with respect to both the trend and changes in the level after the date of entry into force of the agreement. With t_f denoting the date of entry into force of the investigated bilateral trade agreement, we investigate the interval of the 5 years before the trade agreement $\mathcal{I}_p = [TI_{t_f-5}, .., TI_{t_f-1}]$ and 5 years after the agreement $\mathcal{I}_s = [TI_{t_f+1}, .., TI_{t_f+5}]$. We consider an interval of 5 years to be reasonable, as we assume that immediate effects of adaptation processes triggered by the agreement have impacted the topology on the ITN within this period. Furthermore, the investigation of larger time intervals would increase the risk that external factors that are not attributed to the agreement will influence the TI as well. On the other hand, shorter time intervals would make it more difficult to decide if the trade agreement had a lasting effect on the trade relations.

We define the z-score

$$z := \frac{\mathrm{E}(\mathcal{I}_s) - \mathrm{E}(\mathcal{I}_p)}{\sqrt{\mathrm{Var}(\mathcal{I}_p)}} \,, \tag{4.2}$$

with $\mathrm{E}(\mathcal{I}_\bullet)$ and $\mathrm{Var}(\mathcal{I}_\bullet)$ representing the mean value and the variance within the interval, respectively. This z-score sets the values of *TI* after t_f in relation with previous levels of the variable. If the variance is large, the difference of the mean values in the two intervals are less significant and the z-score is reduced.

Next to changes in the level of the TI, we are interested in the trend after the date of entry into force of an agreement. Thus, we perform a linear regression on the interval $[I_{t_f}, .., I_{t_f+5}]$. To recognize if there is an oscillating or saturating behavior of the time series, we additionally perform a segmented linear regression [100]. The model of segmented linear regression can be written as

$$y_s(t) = \gamma_0 + \gamma_1 t + \gamma_2 (t - \psi)\theta(t - \psi) + \epsilon \,, \tag{4.3}$$

with the Heaviside function θ, the break-point ψ, parameters γ_i and an assumed Gaussian error term ϵ of the segmented regression model. In contrast to linear regression, the model in equation (4.3) allows for the identification of one local extreme value in the investigated interval. This would be represented in a change of signs of the slopes between the two segments. However, the segmented model introduces two degrees of freedom compared to linear regression. To prevent overfitting, we select a model by utilizing the Akaike Information Criterion [101] and choose between the segmented or non-segmented model by selecting the model with the lower AIC-score:

$$AIC = 2d - 2\ln\hat{L}\ . \tag{4.4}$$

Here, d denotes the degrees of freedom and $\hat{L}$ the maximum likelihood estimator of the model. If the segmented model is selected, we categorize the trend of TI after t_f according to the sign of the slopes in the respective segments of the fit: $++$ $(--)$ for a continuous positive (negative) trend, whereas there is no continuous trend when the sign of slope changes as in $+-$ and $-+$. If the AIC selects the linear model, we further analyze the relevance of the slope and compare the estimated standard deviation of the error $\hat{\sigma}_\epsilon$ in the regression model y_l with the difference $\Delta\hat{y}_l := |\hat{y}_l(t_f) - \hat{y}_l(t_f + 5)|$. If $\hat{\sigma}_\epsilon > \Delta\hat{y}_l$ we do not consider the estimated slope of the linear model to be relevant and categorize the trend as o. Otherwise, we denote the sign of the slope as the trend's category.

Combining the properties of the trend and changes in the level of the TI time series, we define the impact index Π of the bilateral trade agreement as

$$\Pi(\mathcal{G}'_x, \mathcal{G}'_y) := \begin{cases} 1 & \text{if } z > 1 \text{ AND } (+\mid ++) \\ 0.5 & \text{if } 0 < z < 1 \text{ AND } (+\mid ++) \\ -0.5 & \text{if } -1 < z < 0 \text{ AND } (-\mid --) \\ -1 & \text{if } -1 < z \text{ AND } (-\mid --) \\ 0 & \text{else} \end{cases}\ . \tag{4.5}$$

The definition of the *BTA impact index* Π in equation (4.5) is a conservative measure in the sense that it prevents false classifications of impacts. These might arise in scenarios of recoveries to previous levels of TI in the case of a sharp drop or peak at around the time of the date of entry into force of the bilateral trade agreement. In the case of a recovery to previous levels after a sharp drop at t_f, the positive trend and the negative z-score would result in a value $\Pi = 0$. This is reasonable, as the rising trend is then possibly attributed to a recovery to previous level instead of to the impact of the bilateral trade agreement. As the TI most commonly do not follow a Gaussian distribution, we utilize the coarse classification of equation (4.5) and do not consider the precise value of z in the definition of Π. In general, a more sophisticated approach to assess potential changes in the level of a random variable would include an analysis of variance (ANOVA), most likely via the Mann-Whitney U test. However, the small sample size of TI prevents a meaningful interpretation of the

test results in this case, which is why we refrain from performing explicit statistical significance testing at this point.

As for *TI*, we speak of $\Pi^{out}(\mathcal{G}'_x, \mathcal{G}'_y)$ as the output BTA impact index of $\mathcal{G}'_x$ to $\mathcal{G}'_y$ and of $\Pi^{in}(\mathcal{G}'_x, \mathcal{G}'_y)$ as the input BTA impact index of $\mathcal{G}'_y$ to $\mathcal{G}'_x$. With a country x's average BTA impact indices Π^{out}_x and Π^{in}_x, we refer to the averages

$$\Pi^{out}_x := E_y[\Pi^{out}(\mathcal{G}'_x, \mathcal{G}'_y)] \quad \text{and} \quad \Pi^{in}_x := E_y[\Pi^{in}(\mathcal{G}'_y, \mathcal{G}'_x)] \tag{4.6}$$

for the export linkages and for the import linkages, respectively. Here, the average is taken over all countries y that have negotiated a trade agreement with x.

4.3. Impacts of bilateral trade agreements (BTA)

We speak of a bilateral trade agreement if there are only two parties involved in the negotiation process. This includes BTAs in which one party is a regional trade block itself, i.e. if the EU negotiates an agreement with Mexico. To analyze the TI between the two partners in these cases, we first aggregate all trade flows within the regional trade block while maintaining the homogeneous sectoral structure. Then the TI is calculated as in equation (4.1), considering the 26 industry sectors of the trade block plus its final demand in the subset of nodes $\mathcal{V}'_x$ and $\mathcal{V}'_y$. The obtained TI is then attributed to all countries that are contained in the regional trade block. As data for the construction of the ITN is available for the years from 1990 to 2013, we can evaluate the BTA impact index of the agreements with a date of entry into force between 1995 and 2008. The World Trade Organization lists 107 of such bilateral trade agreements that we analyze in this study [102].

An overview of all analyzed bilateral trade agreements is given in Table B.1 of the appendix. The probability distribution of the BTA impact index Π^{out} (Π^{in}) under the condition that a BTA has been negotiated are obtained from all pairs of countries $(\mathcal{G}'_x, \mathcal{G}'_y)$ with a BTA in the considered time period (i.e. all pairs in Table B.1) and are illustrated by the solid blue (orange) bars in Figure 4.6. Note that each distribution in solid colors of Π^{out} and Π^{in} has 214 entries, as $\Pi^{\bullet}(\mathcal{G}'_x, \mathcal{G}'_y) \neq \Pi^{\bullet}(\mathcal{G}'_y, \mathcal{G}'_x)$. As explained in section 4.1, economic theory suggests that bilateral trade agreements foster trade activities among the partners. This would result in a stronger TI and thus a positive BTA impact index between the involved countries. A positive BTA index reflects that trade between the involved countries gained in relative importance for a partner in the first years after the implementation of the BTA. Accordingly, we consider a BTA to be effective if its BTA index is positive, i.e. $\Pi > 0$.

To better understand the impact of the BTAs, we investigate the corresponding distributions for pairs of countries that have not negotiated a trade agreement until 2014. For that purpose, we identify the pairs of countries that are not listed in the WTO's Regional Trade Agreement Information System. For these 15199 pairs we calculate the BTA impact index taking 2002 as an arbitrary reference year. The resulting distributions are presented in light colors of Figure 4.6. We observe that the existence of trade agreements increases the likelihood of a positive BTA impact

index. Furthermore, we see that positive BTA impact indices are more common than negative values even for countries without agreement. This represents the fact that the positive trend for international trade in the globalization process affects also trade between countries that have not negotiated any dedicated trade agreement with each other.

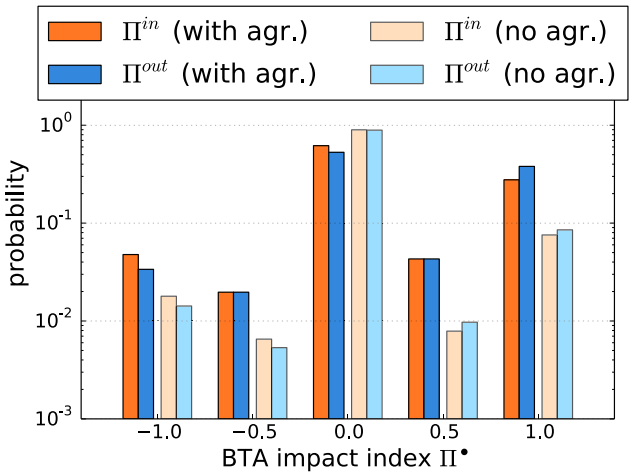


Figure 4.6.: Distribution of the BTA impact indices Π^{in} (orange) and Π^{out} (blue) for all 107 bilateral trade agreements that have come into force between 1995 and 2008 (solid colors). The light colors show the respective results for all pairs of countries that have not negotiated an agreement. The distributions depict the probability of $\Pi^{\bullet}$ for the respective sample of country pairs.

For countries with a trade agreement, we find that most countries x show a positive average BTA impact index in both their export (Π_x^{out}, Figure 4.7A) and import linkages (Π_x^{in}, Figure 4.7B) to their partners. For some countries, e.g. the USA, Australia, India and Columbia, the relative importance of the export linkages to their partners increase to a higher extent than the relative importance of the import linkages from their partners. On the other hand, for other countries, such as the Philippines, Algeria, the southern African countries and Uruguay, the import linkages from their partners gain in importance to a larger extent than the export linkages. Some notable exceptions that do exhibit non-positive values of Π^{out} and Π^{in} for both exports and imports are Ukraine, Bahrain, Jordan, Belarus. The only G20 members with a non-positive value are Indonesia and China.

Thus, China is one of the world's leading trading power that did not increase the relative importance of its agreement partners for its domestic production. By examining the composition of its average BTA impact index in detail in Figure 4.8, we observe that for most of the agreements the level of both input- and output TI of China to its partners has increased after the date of entry into force. China has implemented agreements with Chile (CHL), Hong Kong (HKG), New Zealand (NZL), Pakistan (PAK) and the Association of Southeast Asian Nations (AS) within the investigated

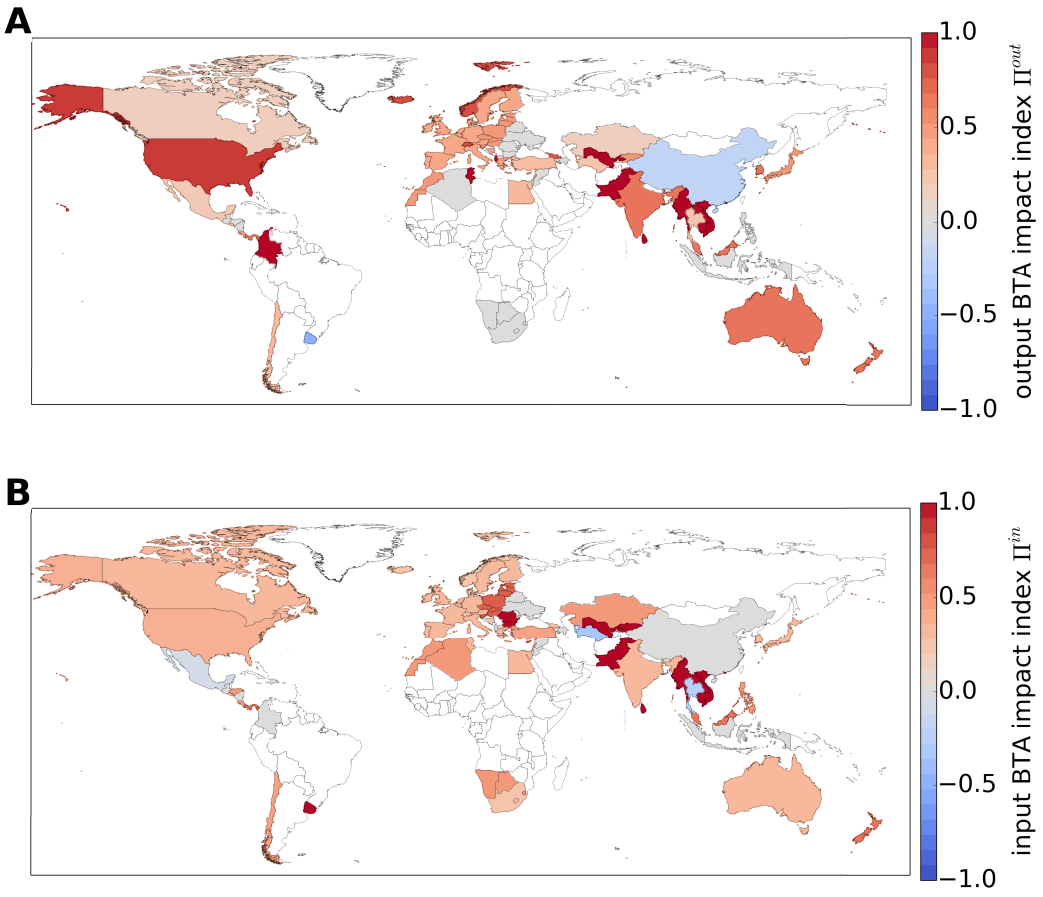


Figure 4.7.: World maps of the average BTA impact indices for the countries' export linkages (Π_x^{out}, A) and for the countries' import linkages (Π_x^{in}, B). Red values indicate that on average the relative importance of the partners have increased for the country. The average is taken over all the trade partners to which the country has negotiated a bilateral trade agreement between 1995 and 2008.

time period. A continuing positive trend cannot be observed for any agreement. This implies that China's agreement partners show no lasting increase in importance for China's production after the trade agreement. Considering the TI of the partners to China, the relative importance of the partners' export and import linkages to China continuously increased within the first 5 years of the implementation period of the agreement. These results imply that China's importance for the partners' economies has increased. On the other hand, we have seen above that the economies of China's agreement partners have not become more important for China's economy after the BTAs have come into force. This finding supports the hypothesis that China's motivations in the negotiation processes are more likely driven by the objective to acquire more influence on the partners than by the objective to open new markets for the Chinese economy.

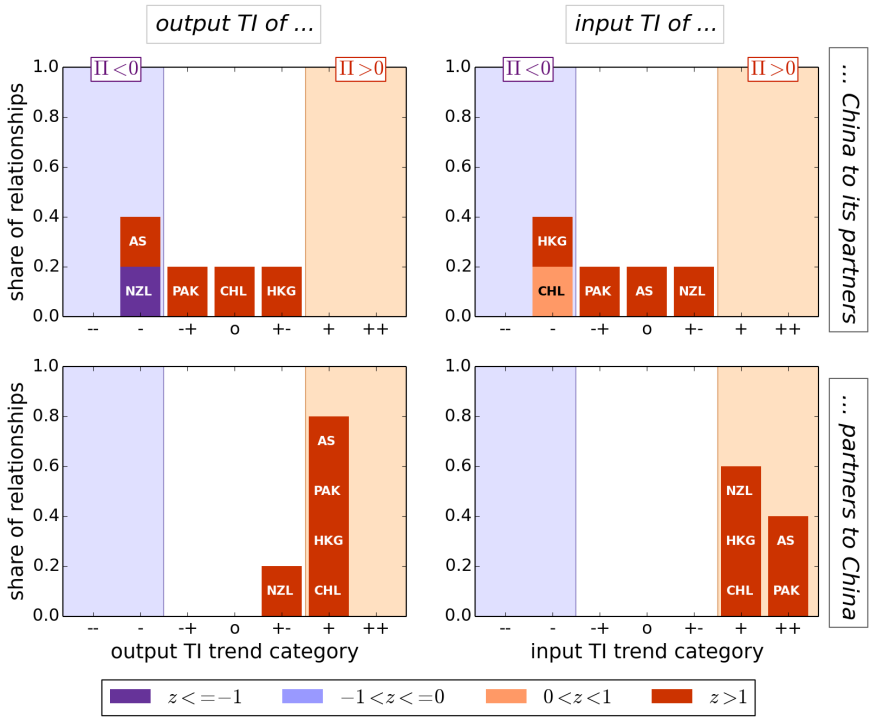


Figure 4.8.: Results of the analysis of the TI of China to her partners (top panels) and of the TI of the partners to China (bottom panels) in the 5-year period after the date of entry into force t_f. The left (right) panels describe the results for the export (import) linkages of the investigated country. The partners of China with a bilateral trade agreement are Chile (CHL), Hong Kong (HKG), New Zealand (NZL), Pakistan (PAK) and the Association of Southeast Asian Nations (AS). A red (violet) z-score implies a higher (lower) level of TI in the 5-year period after t_f. Trade agreements in the red (violet) shaded regions indicate a positive (negative) trend of the TI between the partners after t_f. A positive (negative) value of BTA impact index Π is assigned to a trade agreement if it is succeeded by both a positive trend and a higher level (negative trend and a lower level) of the TI.

A different picture emerges for the bilateral trade agreements of the USA (see Figure 4.9). The USA has negotiated 8 bilateral agreements in the considered time period, namely with Australia (AUS), Bahrain (BHR), Chile (CHL), Dominican Republic (DOM), Jordan (JOR), Morocco (MAR), Singapore (SGP) and the Central American Common Market (CA). All partners of the USA have become more important for US exports with the exception of Jordan. While the importance of the partners' imports for the USA reaches a higher level after the trade agreement, only Australia, CA and Jordan show an enduring positive trend in the input TI of the USA. In the opposite direction, the US have not become relatively more important for their trade agreement partners. This holds for both the export and import linkages of the partners to the USA (except for the import linkages of Morocco). In some cases, even a value of $\Pi < 0$ with both negative trends and lower levels after the agreement of the TI of the partners to the USA are observed. From these findings, we can conclude that the US has opened new markets for their economy to increase trade with their partners above average. However, the partners' economies have not become more dependent on the US after the agreements have come into force. Therefore, these results support the hypothesis that the motivations of the US in BTA negotiations are more market driven compared to the Chinese trade policies.

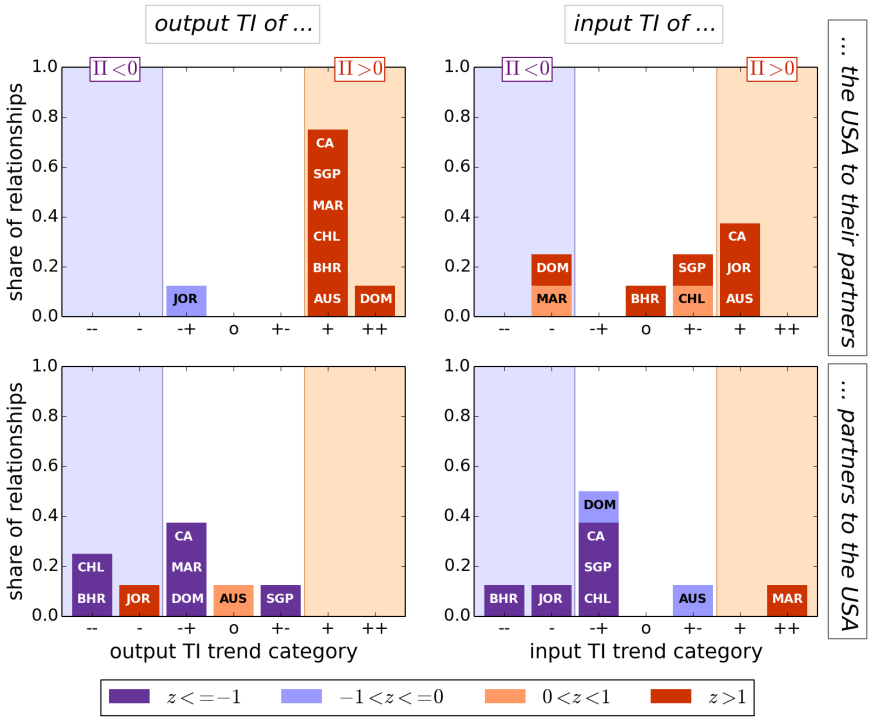


Figure 4.9.: The trade profile for the USA, with definitions as in Figure 4.8. The USA has negotiated bilateral trade agreements with Australia (AUS), Bahrain (BHR), Chile (CHL), Dominican Republic (DOM), Jordan (JOR), Morocco (MAR), Singapore (SGP) and the Central American Common Market (CA) within 1995 to 2008.

4.4. Discussion

We provide quantitative evidence that BTAs mostly result in a stronger trade interconnectedness between the involved partners in the first years after the agreement came into force. This strengthens the conclusions from previous studies on direct trade flows in the gravity model, as we take indirect effects that are mediated by cross-sectoral input-output linkages into account as well. A positive BTA impact index can hint towards the suggestion that one main objective in the negotiation process for these countries consisted of obtaining easier access to the partner's market. The cause of a positive BTA index can be twofold: On the one hand, with decreasing trade barriers new market opportunities can be unlocked to result in an above average increase in trade between the agreement partners. On the other hand, a positive impact index is also achieved if existing trade relations are substituted in favor of the agreement partner and at the expense of a third party. We find also BTAs with a non-positive BTA impact index. In these cases the hypothesis that a trade agreement negatively affects third parties is unlikely. From countries with a negative BTA impact index, we can conclude that they either pursue a different strategy than to boost the bilateral trade with the partners or that they did not achieve their aims in the negotiation process of the agreement.

We present that most economies of the western world as well as Japan and South Korea increase the interconnectedness to their partners for their export linkages to a larger extent compared to their import linkages. This can indicate that these economies mainly focus in the negotiations on developing new sales markets for their domestic economy. On the other hand, countries in southern Africa, Uruguay and the Philippines have increased the interconnectedness to their partners in import linkages to a larger extent than in their export linkages.

China is the only member of the G20 with a negative impact on the trade interconnectedness for the export linkages to her partners. This can suggest that China pursues a different strategy in the trade agreement negotiations than most other countries. The lack of a persistent increase in the TI of China to its partners can be explained by two predominant factors: On the one hand, China has already enjoyed easy access to the more open markets of its partners, such as in the case of Hong Kong, before the bilateral trade agreement [96]. On the other hand, China has continuously increased both the number of trading partners and her international trade volume since the 1990s. Within these developments, the bilateral trade agreements of China do not trigger a disproportionally high and lasting increase in bilateral trade with her agreement partners compared to China's other trade activities. In contrast, as China becomes economically more important for her partners, an obvious suggestion towards China's motives behind her agreements is to increase her economic and political influence among her trading partners.

In contrast to China's strategy stands the agreement profile of the USA. The consistent positive BTA impact index for the export linkages of the USA to her partners emphasizes the focus on the stimulation of her exports during the agreement's negotiations. Although the importance of import linkages from the agreement partners

have also increased for the US, a lasting trend is only observed for Australia, the Central American Common Market and Jordan. However, these increases in trade volume are less important for the partners of the US compared to their market expansions to other third-party countries. The fact that both the input- and output TI of most partners to the US reaches significantly lower levels after the implementation of the trade agreement indicates that even with increasing bilateral trade volumes to the US, the economies of the partners become less dependent on trade with the US.

With bilateral trade agreements affecting the trade interconnectedness between the involved economies, the question arises if the topology of the ITN provides additional insights into the correlation structure of the industries' economic performance. To address this issue, we continue in the next chapter with the search for empirical evidence towards the hypothesis that industries that are strongly interconnected via trade, or show some other characteristic topological property in the ITN, are also correlated in their value added.

Chapter 5.

World trade and correlations in economic performance

In the previous chapters, we have provided empirical evidence for the considerable growth of international trade during the process of globalization. The supply chains for the production of goods and services have become increasingly complex, spreading often over multiple countries and economic sectors. As we have seen, this process is arguably accelerated by bilateral trade agreements. In this context, an empirical assessment to quantify the role of trade as a mediator of supply and demand spillovers has become increasingly important. Traditionally, direct trade relations between the industries have been regarded as the mediators of these spillovers. It is intuitive to assume, however, that with increasing network density the establishment of new trade relations becomes less costly compared to new links in very sparse networks. Cheaper substitution possibilities might potentially damp the impacts of spillover effects of direct trade relations. Thus, with increasing trade network connectivity, the question arises if higher-order relations become more important in explaining a national sector's susceptibility to supply and demand changes of its trading partner. In this chapter, we investigate empirically to what extent the topological properties of the ITN provide information about the statistical association between the time series of two industry sectors' value added of their respective production. For this purpose, we construct the *World Economic Performance Network (WEPN)* that describes the correlation structure among the nodes' generated value added. We then compare the topologies of the WEPN and the ITN and identify which first-order and higher-order network measures of the ITN can serve as predictors for the existence of a link in the WEPN. The next sections follow in parts closely the presentation in publication P2.

5.1. Introduction

In a globalized economy the decrease of both trade barriers and costs for transportation and communication are considered main drivers for the specialization of companies and countries in particular stages of production. In the economic literature, these processes have been referred to as vertical specialization [103] and fragmentation of production [104]. Economic agents and entities, such as companies and multi-national corporations, assess their outsourcing opportunities on an international level [105]. Numerous studies have been performed to quantify the degree of economic integra-

tion and interconnectedness of world trade [106–108]. These studies conclude that fragmentation of production has increased considerably in the last years, stretching also beyond regional trade blocks [107, 108]. However, as presented in chapter 3, the process of economic integration is yet far from being completed, leaving potential for a further integrated world economy [106].

In this context, we address the question to what extent international trade serves as a transmitter of demand spillovers and economic shocks at the sectoral level. Evidence has been reported that direct trade channels play a significant role in the transmission of a global recession [30]. Furthermore, it has been shown that international fragmentation can have distinct consequences for different players in the economy [109, 110] as different companies are subject to different opportunities and challenges. For example, previous studies conclude that an individual company's offshoring strategy depends on the engineering details of the production process [111].

A variety of models of contagion phenomena have been proposed in order to simulate the global spreading of an economic crises in the ITN. Simple diffusion and epidemic models provide valuable insights into the stability and possible dynamics of a crisis spreading in the ITN [23, 28]. More sophisticated approaches introduce agents to the model that locally optimize their production [26, 27]. In these models, also countries with comparatively small economies have the potential to trigger a global crisis due to the complex network structure. Notably, these results have been obtained based on the assumption that only existing direct trade relations serve as immediate transmission channels of supply and demand changes. While this is arguably an intuitive assumption, there is little corresponding empirical evidence. As the ITN is far from being static, the agents in the network might adapt quickly to a disruption in the network by substituting their input sources and reaching out for new markets. Thus, the importance of direct trade relations might be overestimated in the existing models.

In this chapter, we investigate empirically to what extent local topological features in the ITN are able to indicate the appearance of correlations between the sectoral *value added growth (VAG)* of two industries. The VAG serves as a good indicator for economic productivity and allows for a comparison between industries of different size. We analyze both positive and negative correlations in the VAG of the nodes. A positive correlation indicates that the nodes follow the same business trend while a negative correlation could be expected among two strongly competing nodes, for example. For the estimation of the correlation between two industries' VAG, we put each industry's yearly economic performance in the global context and define the WEPN by identifying node pairs with significantly correlated VAG. We then investigate the pairs' relationship with respect to their local topological properties in the ITN to determine if direct trade relations (first-order) or higher-order network measures are suitable predictors for the existence of a link in the WEPN.

The remainder of this chapter is organized as follows: In section 5.2 we describe the construction of the WEPN from the data in detail. We then analyze the topology of the WEPN for both positive and negative correlations in section 5.3. Section 5.4

compares the relation between the ITN and the WEPN and we conclude this chapter with a discussion in section 5.5.

5.2. World Economic Performance Network (WEPN) construction

We construct the ITN with an inflationary adjusted threshold as described in section 3.2 from the Eora MRIO database [32, 33]. The WEPN consists of the identical set of nodes $\mathcal{V}$ as the ITN in which each node v_i is attributed with an industrial sector and the associated country. To assess an industry's economic performance and to construct the links in the WEPN, we analyze the growth rate of the value added from the **VA** block of the Eora data. To account for distortions arising from individual national inflation and exchange rate fluctuations, we convert the time series of the value added from current to real US $ by utilizing annual exchange rates and deflationary indices provided by the World Bank [112, 113]. The conversion from current dollars US_t in year t to real dollars US_{t_b} with a fixed reference year t_b is

$$US_{t_b} = \frac{LCU_{t_b}}{e_{t_b}} = \frac{LCU_t}{g_t \cdot e_{t_b}} = \frac{US_t \cdot e_t}{g_t \cdot e_{t_b}} \,, \tag{5.1}$$

where LCU_t denotes the local currency unit, $g_t = LCU_t/LCU_{t_b}$ is the GDP deflator of the country and $e_t = LCU_t/US_t$ is the exchange rate in year t. Reliable data required for the conversion is available for 126 countries, as listed in Table A.2 of the appendix. We therefore restrict our analysis to these countries.

The industry's value added serves as a better indicator for economic performance than an industry's gross output, as it accounts for how well an industry adapts to sudden supply or demand shocks. An industry can adapt to these shocks by either redistributing its allocations or by improving its productivity [114, 115]. An analysis of the growth rate instead of the absolute value accounts for the effect of different sizes of economic sectors and national economies. Thus, a comparison of the growth rate of the value added between two industries in economies of different magnitude still offers meaningful economic interpretations.

We note that world trade and the global economic system are far from being stationary. Furthermore, each industry is potentially subject to economic developments at a global scale, such as global economic trends and crises. Therefore, a correction for the global trends is necessary for the recognition of pairwise correlations between industries' performances. We tackle this issue by putting the value added growth

$$\nu a_i^t := (VA_i^t - VA_i^{t-1})/VA_i^{t-1} \tag{5.2}$$

of industry v_i in year t (with the annual value added VA_i^t) in relation to the current global distribution of growth rates. Figure 5.1 shows the distributions of νa_i for the $26 \cdot 126$ analyzed industries for each year between 1991 and 2011. In accordance with

previous studies [116, 117], we see from the assymmetric distributions in the box plot that the growth rates do not follow Gaussian distributions. Regarding temporal changes, we observe that the data reveal large variations in VAG between 1991 and 1993. Furthermore, a decline of the global median is observed in 1998 - 2002. In these years, the global economy experienced particularly unstable conditions, comprising the Asian crisis in 1997 [118], the Russian financial crisis in 1998 [119], the early recession in the developed countries in the early 2000s [120], the burst of the dot-com bubble in 2001 [121] and the Argentinian financial collapse in 2001-2002 [122]. The Great Recession of world markets, triggered by the financial crisis, is also well represented with a negative median of νa in 2009.

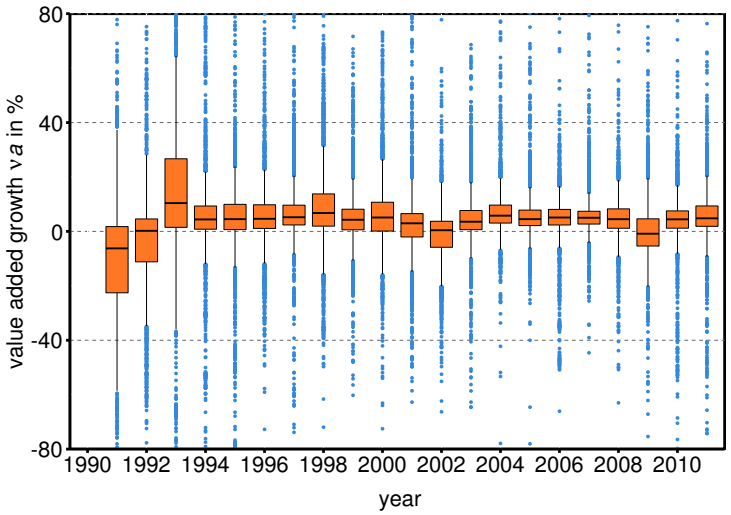


Figure 5.1.: Box plot depicting the median M (black line), lower and upper quartile (hinges of the orange box) of the yearly global distribution of VAG rates. Outliers that exceed the range of $1.5 \cdot \mathrm{IQR}$ (inter-quartile range) from the hinges are plotted as points. In each year, the distribution consists of 3276 national economic sectors (26 industries for each of 126 countries) of the ITN.

To account for these global trends, we obtain the trend-adjusted growth rates g_i^t for node v_i by

$$g_i^t = \frac{\nu a_i^t - M^t(\nu a)}{|M^t(\nu a) - QR_{\pm}^t(\nu a)|} \,, \tag{5.3}$$

where M^t and $QR_{\pm}^t$ denote the median and upper $(+)$ and lower $(-)$ quartiles of the global growth distribution over all national economic sectors in year t, respectively. The upper quartile enters the equation, if νa_i^t exceeds the median. Otherwise, the lower quartile is considered.

With the corrected VAG time series (5.3), we construct a functional network in which the links between the industries are defined by the existence of significant correlations in their VAG. In the following, we refer to VAG to the corrected value

added growth g_i, unless explicitly stated. Hence, this functional network describes and visualizes the interdependencies between the performances of national sectors in the global economy. To identify relevant transmission channels of supply and demand changes in the network, we first focus on the positive correlations. Thus, we define a functional network of positively correlated sectors in terms of their VAG, which we refer to as the *World Economic Performance Network (WEPN).*

The estimation of statistical interdependencies from short time series is challenging [123–125]. For small sample sizes, the Pearson correlation coefficient has several drawbacks: With a high sensitivity for outliers, it provides in general a biased estimate of the true correlation coefficient with a comparatively large standard deviation of the coefficient's sampling distribution [126]. Here, we measure the statistical association $\mathrm{Corr}_{ij} := \mathrm{Corr}(g_i, g_j)$ between the time series of the VAG of sector v_i and v_j using Spearman's rank correlation coefficient [127]. To test for the significance of the obtained values, we compare each value of Corr_{ij} with the distribution f obtained from an ensemble of randomly permuted time series $\{\pi[g_j]\}$. From the distribution $f(\mathrm{Corr})$ we estimate the p-value for each observed Corr_{ij}. By choosing a desired significance level α and obtaining the threshold $\mathrm{Corr}^{\alpha}_{ij}$ with

$$\int_{-1}^{\mathrm{Corr}^{\alpha}_{ij}} f\left(\mathrm{Corr}(g_i, \{\pi[g_j]\})\right)\, \mathrm{dCorr}(g_i, \{\pi[g_j]\}) = 1 - \alpha\,, \tag{5.4}$$

we decide if the VAG g_i and g_j of two sectors show a significant positive correlation for the years between 1991 and 2011. If $\mathrm{Corr}_{ij} > \mathrm{Corr}^{\alpha}_{ij}$, we consider the economic performance of industry v_i and v_j to be positively correlated or linked to each other. In this case we draw a directed link from v_i to v_j. We perform this procedure described above for each pair (i, j). The directed nature of the WEPN arises from the fact that the significance test of the statistical association described in equation (5.4) is not symmetric with respect to the exchange of g_i and g_j.

To additionally assess the robustness of the construction procedure of the WEPN, we create alternative WEPNs by choosing different surrogate methods for estimating the significance of the pairwise correlations. Thus, instead of randomly permuting the time series g_i, we estimated $\mathrm{Corr}^{\alpha}_{ij}$ in equation (5.4) with Fourier-based surrogates [128] and with surrogates created by the block bootstrapping method. The Fourier-based surrogates have been created by multiplying random phases to the Fourier transform of g_i and then transforming the series back to the time domain by the inverse Fourier transform. In the block bootstrapping method, we randomly permute blocks with a length of τ years from the original time series instead of single points to create the surrogate time series. To decide upon a suitable value for τ we analyze the distributions of autocorrelations for the VAG series g_i at different lags. We estimate the autocorrelation ACF of one time series g_i as

$$\mathrm{ACF}(\tau) = \frac{1}{n \cdot \mathrm{Var}(g_i)} \sum_{t=\tau}^{n} (g_i^t - \mathrm{E}(g_i))(g_i^{t-\tau} - \mathrm{E}(g_i))\,. \tag{5.5}$$

We observe that at a lag of 3 years, the autocorrelations are sufficiently small, with an inter-quartile range < 0.2 (see Figure 5.2). We compare the resulting WEPNs by measuring the fraction of pairs where a link exists in one network and is absent in the other. At a significance level of $\alpha = 0.05$, this fraction accounts for 1.23 % (1.21 %) when comparing the WEPN from shuffled surrogates with the WEPN from block bootstrapping surrogates (Fourier-based surrogates). The results presented in the following are qualitatively robust with respect to the choice of surrogates.

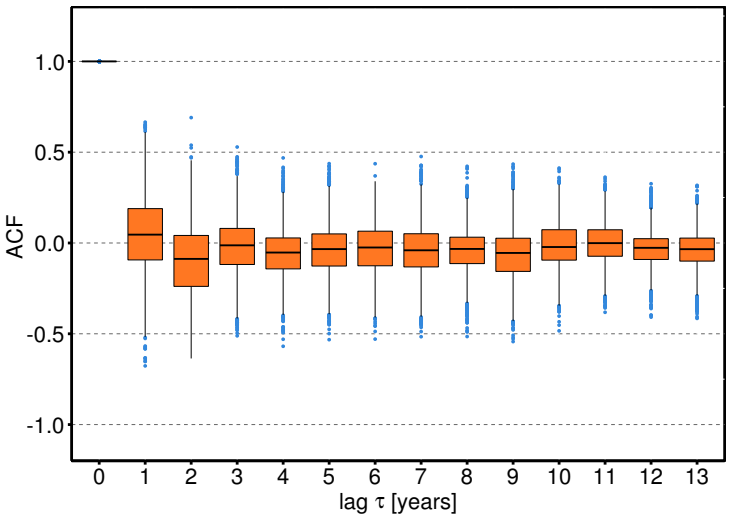


Figure 5.2.: Distributions of the autocorrelation (ACF) of the VAG time series g_i as a function of the time lag τ. The definitions of boxes, points and whiskers are defined as in Figure 5.1.

Next to the construction of the WEPN for positively correlated sectors in VAG, we investigate the case of sectors that are competing for identical groups of consumers. For this purpose, we analyze the significant negative correlations in VAG. For example, a negative correlation between sectors could be expected, if one sector experiences production losses caused by a natural disaster or a national economic crisis, and another sector steps in as a substitute [43]. The construction of the functional network with negative correlations follows the methodology described above. For negative correlations, we perform a left-sided significance test by obtaining the threshold $\mathrm{Corr}_{ij}^{\beta}$ with

$$\int_{\mathrm{Corr}_{ij}^{\beta}}^{1} f\left(\mathrm{Corr}(g_i, \{\pi[g_j]\})\right) \mathrm{dCorr}(g_i, \{\pi[g_j]\}) = 1 - \beta\,, \tag{5.6}$$

for a significance level β. Again, $f(\mathrm{Corr})$ is the distribution of the Spearman correlation of the surrogates for each observed value Corr_{ij}. A link between two industries (i, j) in the network of negative correlations is drawn if $\mathrm{Corr}_{ij} < \mathrm{Corr}_{ij}^{\beta}$.

5.3. Topology of the WEPN

5.3.1. Positive correlations

Before we discuss the relationship between the topologies of the ITN and the WEPN in the next section, we provide an overview of the topology of the WEPN first. In particular, we investigate if the national or sectoral attribution of nodes provide already information about the presence or absence of positive correlations in VAG. For this purpose, we investigate the link density ρ of the network and the unweighted modularity in directed networks Q_d (see chapter 2) of the national partition $\mathcal{C}_c$ and sectoral partition $\mathcal{C}_s$ of the WEPN. The definition of these partitions is identical to the definition of the partitions defined in the ITN (cf. section 3.3). In the WEPN, the link density ρ is interpreted as the fraction of significantly correlated pairs of sectors. Figure 5.3A illustrates the link density for both internal and cross-links of the partitions $\mathcal{C}_c$ and $\mathcal{C}_s$, respectively. The high link density of 70 % (for $\alpha = 0.05$) for all internal links within one subnetwork of the national partition confirms the well known fact that despite globalization, industries within the same country are still significantly stronger connected among each other than with their international partners [106].

For the full network, we observe a much lower link density of 6.7 % at significance level $\alpha = 0.05$. The link density is only slightly larger for pairs that belong to the same industrial sector in $\mathcal{C}_s$ (7.4 %). For more restrictive significance levels, the link density for internal links in $\mathcal{C}_s$ is even lower compared to the full network. These results further support the persisting division of the world economy into national partitions, as industries within the same country also show the most significant correlations in VAG. Due to the specific characteristics in the subnetworks and since the difference $\rho - \alpha$ is positive at all significance levels α, we conclude that the WEPN captures real structural information on economic performance interdependencies beyond the expected rate of α for the false positives.

In the context of the WEPN, the reciprocity r describes the robustness of the construction procedure described in equation (5.4) with respect to the exchange of the time series g_i and g_j for each pair of sectors (i, j). As shown in Figure 5.3B the reciprocity is particularly high for pairs of sectors within one country. For the full network we observe $r = 97.5$ % at $\alpha = 0.05$. With more restrictive significance levels, the reciprocity r decreases in all partitions.

Further important insight into the topological structure of the WEPN is obtained by analyzing its community structure which is defined through a community detection algorithm that seeks to maximize a partition's modularity [74]. Here, we analyze the modularity Q_d of partitions in the WEPN: (i) the national partition $\mathcal{C}_c$ that groups together industries from the same country, (ii) the sectoral partition $\mathcal{C}_s$ and (iii) the partition $\mathcal{C}_m$. The modularity scores of these partitions are shown in Figure 5.3C. Remarkably, the national partition $\mathcal{C}_c$ does not show a strong modularity at $\alpha = 0.05$. However, the community detection finds a partition $\mathcal{C}_m$ with high modularity scores, indicating many significant positive international correlations in

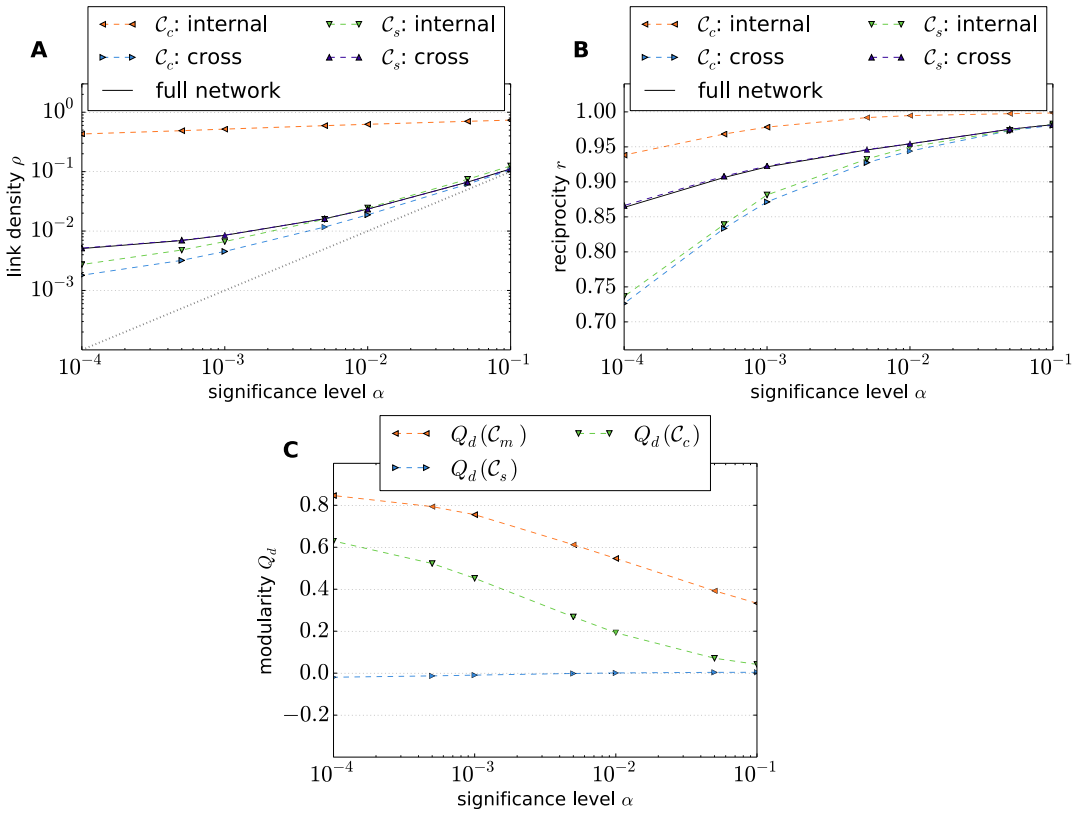


Figure 5.3.: Link density (A) and reciprocity (B) in different subnetworks, and modularity values (C) for different partitions of the World Economic Performance Network (WEPN) in dependence of the chosen significance level α. The gray dotted line in (A) depicts the expected amount of false positives from the significance test.

VAG. With more restrictive significance levels, the national partition becomes more pronounced, yet strong positive international correlations also remain intact. The sectoral partition, however, does not resemble a partition with a community structure in the network-theoretical sense at all significance levels.

5.3.2. Negative correlations

We now compare the topological properties of the WEPN described above with the respective characteristics of the functional network obtained from negative correlations. In Figure 5.4A the link density ρ at a significance level of $\beta = 5$ % accounts for 0.057. In contrast to the WEPN, the link density in the national partition $\mathcal{C}_c$ takes low values that are smaller than the expected false positive rate for most significance levels. This observation harmonizes with the previously presented results: industries of the same country are very unlikely to show negative correlations, as already 70 % of the pairs have significant positive correlations at a 5 % significance level. For international links and in other partitions the link densities are still lower than in the WEPN at all significance levels. However, the numbers of significant negative correlations exceed those that are expected from the false positive rate of the significance test. The reciprocity in Figure 5.4B shows in general a lower value than in the WEPN, except for the internal links in $\mathcal{C}_c$. These links, however, show low statistics and are therefore subject to larger fluctuations, leading to the non-monotonous curve for these links. With respect to the community structure, both the national and sectoral partitions in this network do not show a community property in the sense of complex network theory, with values of the modularity close to zero for all significance levels. On the other hand, using the community detection algorithm $\mathcal{C}_m$ a modular structure in the network can be obtained for low values of β. From these findings we conclude that despite a comparably small link density, the functional network from negative correlations in VAG exhibits some inherent structure that is neither explained by the national nor the sectoral partition.

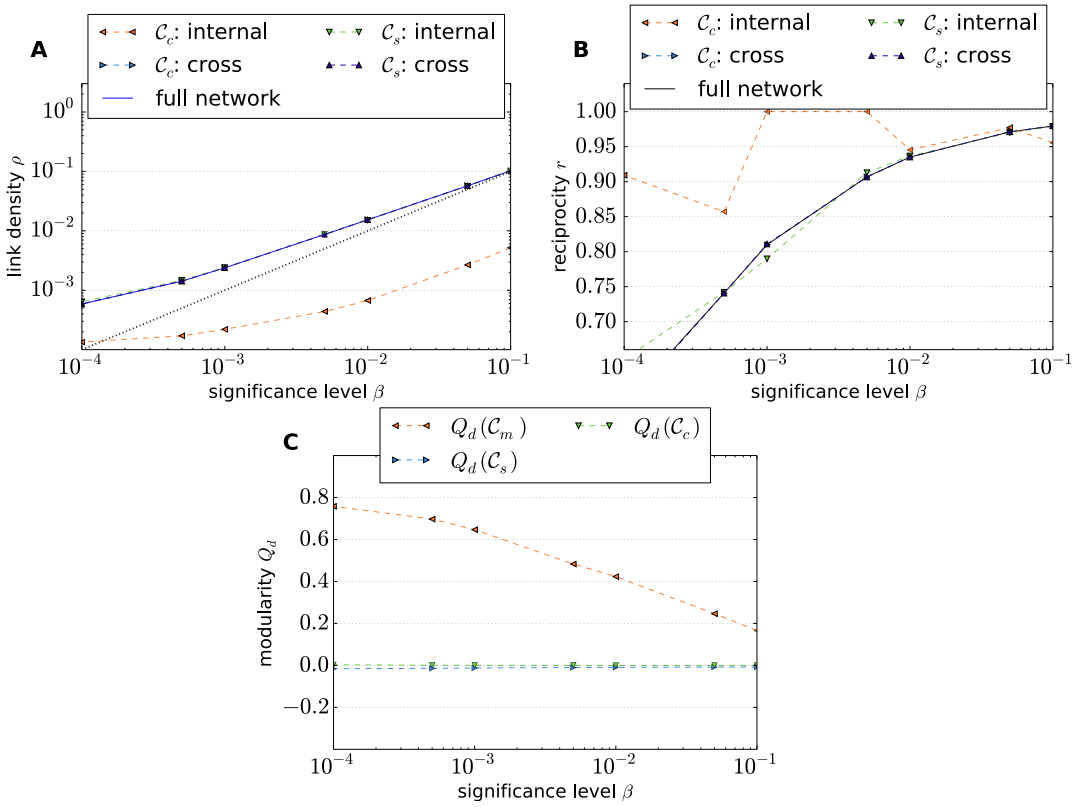


Figure 5.4.: Link density ρ (A), reciprocity r (B) and the modularity Q_d (C) for different partitions of the network constructed from negative correlations in dependence of the chosen significance level β. The black dashed line in (A) denotes the expected background of false positive links from the significance test.

5.4. Relations between the ITN and the WEPN

5.4.1. WEPN from positive correlations in VAG

First-order trade relationships

We have seen in section 5.3.1 that industry pairs in the national partition are likely to be significantly positively correlated in the VAG. Nevertheless, there also exist relevant correlations among industry sectors from different countries. In order to obtain a better understanding of the structural indicators for these correlations, we compare the link structure of the WEPN with the ITN. As assumed by most spreading models of economic crisis, direct trade relationships would suggest a corresponding dependence of VAG between two industries [26, 28]. Since the links in the ITN can be regarded as an aggregation over all supply chains, changes in the economic performance of a supplying industry will affect the linked consuming industries by changes in the availability of production factors. In the case of an economic crisis that affects the supplying industry, the output levels of the supplier are likely to decline. Thus, input costs for the industrial consumer are expected to increase, either due to the need of substitution or due to increasing prices of the input factors. This mechanism also works analogously in the opposite direction: a decline in demand might transmit to the supplying industry as the latter is confronted with decreasing purchasing power on the market.

These arguments suggest that the direct trade volume between two industries might serve as an indicator for the likelihood of a link in the WEPN. To analyze empirically if trade relations indicate positive correlations in VAG between industries, we compute the yearly ITNs from the Eora database (covering the years 1990 - 2011, see section 3.2) the temporal mean of the trade volume w_{ij}. We then compare the frequency distributions of w_{ij} for industry pairs with significantly positively correlated (i.e. linked pairs) and not significantly positively correlated VAG (i.e. not linked pairs) in the WEPN. For the construction of the WEPN we choose a confidence level of 95 % that corresponds to $\alpha = 0.05$. With the separation of industry pairs into these categories, we assess the applicability of various topological properties of the ITN as an indicator for a link in the WEPN by quantifying differences of the two distributions of the investigated topological property. The obtained probability distributions for the temporal mean of the trade volume w_{ij} are presented in Figure 5.5A (note the logarithmic scales). The corresponding bar chart that counts the individual entries in each bin is shown in Figure 5.5B. We find that the distribution for industry pairs with no significant positive correlation in their VAG is shifted towards smaller values of the trade volume. This result suggests that industry pairs with a strong trade relation are more likely to be positively correlated in their VAG than industry pairs with a small trade volume.

Next, we investigate how substitution opportunities for both suppliers and consumers provide information towards the link structure in the WEPN. For that purpose, we utilize the output and input dependency measures $p^{\bullet}_{ij}$ (cf. section 2.4) and compare their distributions with respect to their linking properties in the WEPN. The input

dependency p_{ij}^{in} quantifies the share of inputs of industry v_j supplied by industry v_i. On the other hand, the output dependency p_{ij}^{out} describes the ratio of v_i's outputs that is sold to v_j. A high value of the dependency measure can therefore indicate that one industry has low substitution opportunities and is highly dependent on the other.

As for the trade volume w_{ij}, we consider here the temporal mean of the dependency measures over the 22 analyzed years. The probability distributions of the two dependency variables for linked and not linked pairs in the WEPN are shown in Figure 5.5C,E (with the corresponding bar charts in Figure 5.5D,F).

Table 5.1.: Parameters of the definitions of the histograms in Figure 5.5 and Figure 5.7 as described in equation (5.10).

First-order variables	$\mathbf{x_L}$	$\mathbf{x_R}$
trade volume w	-10	-1
input dependency p^{in}	-8	1
output dependency p^{out}	-8	1
Higher-order variables	$\mathbf{x_L}$	$\mathbf{x_R}$
edge betweenness b	-3	5
common neighborhood input dependency	-8	1
common neighborhood output dependency	-8	1

To quantify the difference between the probability distributions of the investigated variable x for linked ($li(x)$) and non-linked ($nl(x)$) pairs of sectors in the WEPN, we compute the Kolmogorov-Smirnov statistic KS [129]

$$KS = \sup |LI(x) - NL(x)| \tag{5.7}$$

$$\tag{5.8}$$

and the Jensen-Shannon divergence JS [130]

$$JS = \frac{1}{2}\left[\sum_l \left(LI_l \ln \frac{2LI_l}{LI_l + NL_l} + NL_l \ln \frac{2NL_l}{LI_l + NL_l}\right)\right] . \tag{5.9}$$

Here, $LI(x)$ and $NL(x)$ describe the cumulative distribution functions associated with $li(x)$ and $nl(x)$, respectively. LI_l and NL_l denote the probability of the l-th bin in the histograms of the corresponding functions. For the practical computation of JS, we define a histogram from the measured probability distribution. The bins are spaced evenly in log scale, with the position x_l of the left edge of bin l defined as

$$\log_{10} x_l = x_L + \frac{l(x_R - x_L)}{99} , \tag{5.10}$$

where x_L and x_R denote the lower and upper limits of the histogram, respectively.

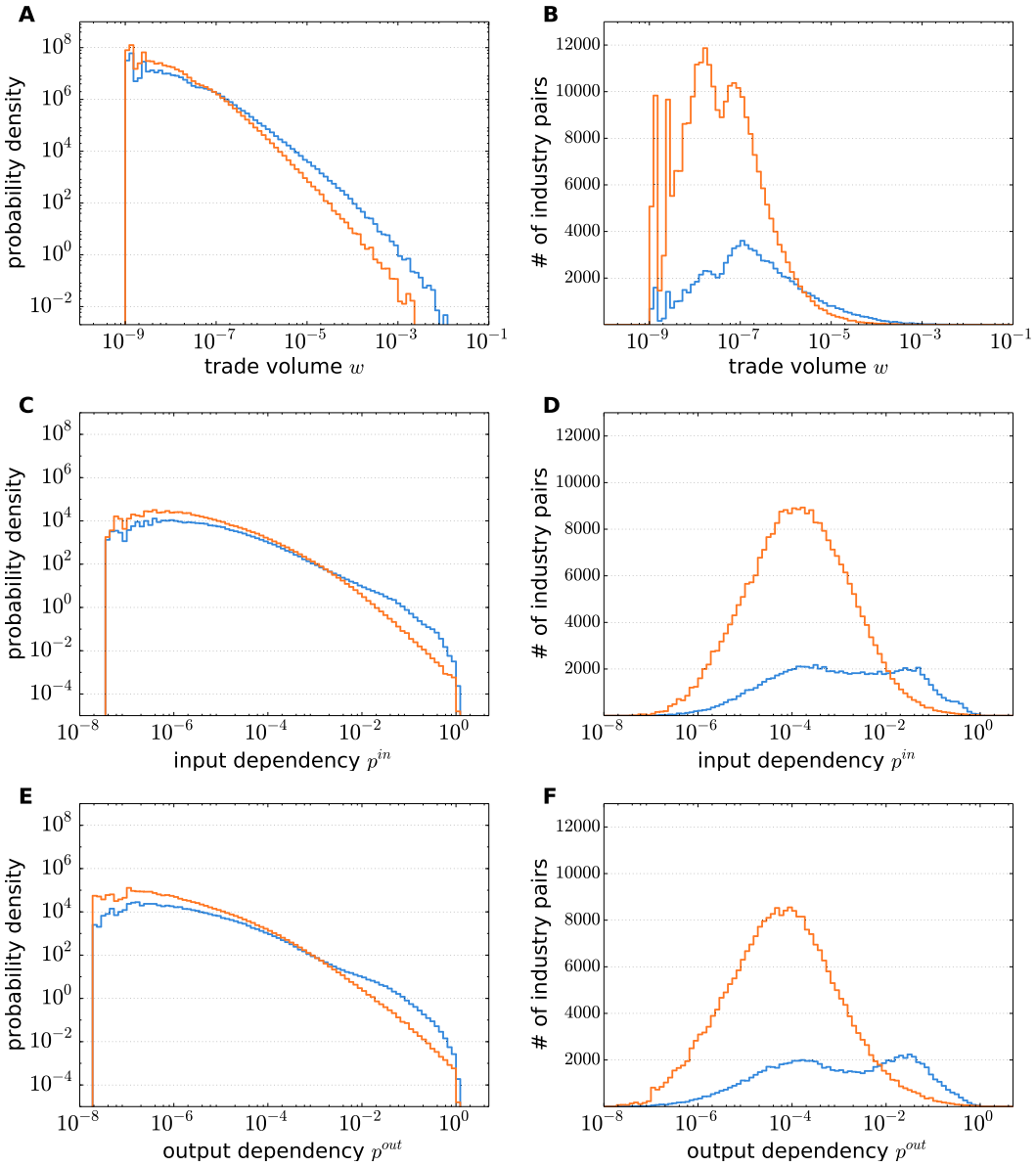


Figure 5.5.: Probability distributions of mean trade volume w_{ij} (A), input dependency p_{ij}^{in} (C) and the output dependency p_{ij}^{out} (E) between all pairs of industry sectors with significantly positively correlated (blue) and not significantly positively correlated (orange) VAG time series at a 0.05 significance level. The corresponding bar charts of the respective variable are shown in (B), (D) and (F).

The chosen ranges for the different variables are defined in Table 5.1. The table also includes the properties of the higher-order variables that are described later in the text.

Table 5.2.: Comparison between the distributions of ITN characteristics for pairs of sectors with significantly positively correlated and not significantly positively correlated VAG. Results are also shown for subsets where only national (international) pairs are considered in the distributions.

Variable	KS	JS	KS	JS	KS	JS
	full network		national		international	
First-order variables						
w_{ij}	0.27	0.057	0.21	0.03	0.14	0.014
p_{ij}^{in}	0.29	0.076	0.07	0.008	0.03	0.003
p_{ij}^{out}	0.33	0.095	0.09	0.008	0.08	0.006
Higher-order variables						
b_{ij}	0.07	0.006	0.11	0.02	0.06	0.009
c_{ij}^{in}	0.19	0.041	0.37	0.10	0.13	0.02
c_{ij}^{out}	0.21	0.037	0.30	0.06	0.13	0.014

The results of the KS and JS statistics are summarized in Table 5.2. In the full network, the higher values of KS and JS for the distributions of $p_{ij}^{\bullet}$ compared to w_{ij} demonstrate that the dependency measures are better indicators of positive correlations in the growth rates of value added as the direct trade relationships. This indicates the importance of substitution opportunities of an industry. We observe from Figure 5.5C and E that for linked sector pairs, the probability density exceeds the density for non-linked pairs for higher values of the dependency measures ($p_{ij}^{\text{in(out)}} > 0.002$). This finding can be understood by recalling that sectors are facing additional market search costs for establishing new trade relations, whereas changes in allocations between existing trade relations are easier to pursue. Therefore, a correlation in VAG is more likely between industry sectors with high dependency. Another remarkable finding lies in the fact that the output dependency is a more suitable indicator for a link in the WEPN than the input dependency. This underlines the specific importance of the consuming industry's role in the ITN. If one industry sells its products to predominantly one consumer, a link in the WEPN between the two industries is more likely. From our findings, we conclude that this effect is more strongly expressed than the opposite effect, i.e. if one industry buys its products from predominantly one supplier.

To gain further insights into the impact of trade structures in the WEPN, we have additionally performed the described analysis with a division into two subsets of pairs with exclusively national and international trade relations, respectively. Thus, we distinguish between internal (corrosponding to national) and cross (corresponding to international) pairs of industry sectors in the national partition $\mathcal{C}_c$ as in section 3.3.1.

Within those subsets, both the trade volume and the dependency measures lose their quality as proxies applicable for discriminating between correlated and uncorrelated pairs in VAG (cf. Table 5.2 as well as Figure 5.6). From these findings, we conclude that national trade relations show in general higher values in the dependency measures and are more likely to be correlated in VAG than international industry pairs.

Correlation in value added growth are expected to emerge due to common business cycles that arise from fluctuations in economic activity. Following Keynesian arguments, these correlations are primarily expected between industries within the same country. The strong interconnectedness of sectors from the same country is represented in both the ITN with a national link density of about 50 % (see section 3.4.3) and the WEPN with a national link density of 70 %. Effects of the interconnectedness between sectors at an international level can be expected by the process of creative destruction [131], i.e. that new technologies from foreign countries might induce structural changes of a domestic economy. However, due to the coarse sectoral resolution of the utilized data, finding any evidence for this process proves to be challenging.

Higher-order trade relationships

Although the previous results show that first-order trade relations between industries are important indicators for positive correlations in industry's performance, they cannot account for the complete correlation structure. Spillover effects, cascading mechanisms and macroeconomic sentiment further contribute to positive correlations between the VAG of industries that have not established a direct trade relationship [132]. This statement is emphasized by the fact that in 87.2 % of all links in the WEPN, no relevant direct trade links between the industries have been established. Furthermore, we have observed that the first-order variables are no good proxies for correlations within the subsets of national and international trade relations.

To better understand how international trade mediates the cascading spillover effects, we investigate the link structure in the WEPN of industry pairs that show higher-order connections in the ITN, i.e. that have not necessarily established a direct trade relation. As input-output tables can be considered as the outcome of an aggregation over all supply chains, the edge betweenness could be a relevant indicator for a trade link's importance in connecting two industries along a longer supply chain. Similar to the node betweenness from section 2.4, the edge betweenness of a link (i, j) from node v_i to node v_j is defined to quantify its importance in connecting any two nodes in the ITN [59]:

$$b_{ij} := \sum_{kl} \frac{\sigma_{kl}[(i,j)]}{\sigma_{kl}} \,, \tag{5.11}$$

where $\sigma_{kl}[(i,j)]$ is the number of shortest paths between node v_k and node v_l including link (i,j) and σ_{kl} is the total number of all shortest paths between these two nodes. Note that for the sake of simplicity, in the definition of equation (5.11) we do not

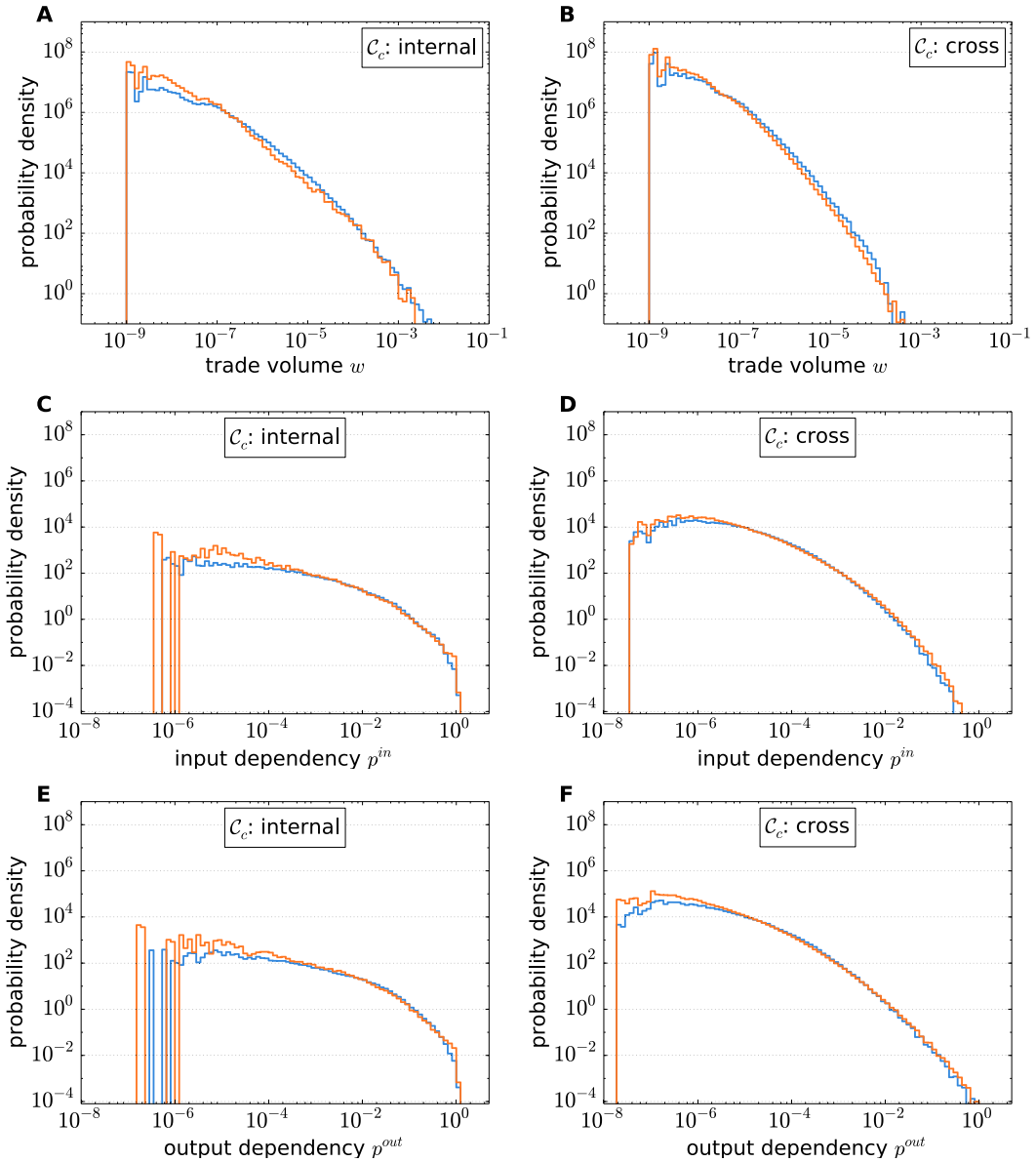


Figure 5.6.: Probability distributions of the trade volume (A,B), the input dependency p_{ij}^{in} (C,D) and the output dependency p_{ij}^{out} (E, F) distinguished between significantly positively correlated pairs (blue) and not significantly positively correlated pairs (orange) at a significance level of $\alpha = 0.05$. The left column depicts the distributions for internal subnetwork pairs in the national partition $\mathcal{C}_c$. The distributions of only international pairs are presented in the right column.

consider weights of the links. An extension of our study to weighted betweenness would be straightforward.

Figure 5.7A reveals that the probability distributions of edge betweenness b_{ij} in the ITN do not exhibit large differences between linked and unlinked industry pairs in the WEPN. This result stresses that due to the high level of aggregation, tracking individual supply chains from input-output tables is practically impossible. Instead, potentially more suitable structural variables of the ITN that may allow to identify positive correlations in VAG are the *common neighborhood input (output) dependency* c_{ij}^{in} (c_{ij}^{out}),

$$c_{ij}^{\text{in}} := \sum_{k \in \mathcal{V}_i' \cap \mathcal{V}_j'} p_{kj}^{\text{in}} \quad \text{and} \quad c_{ji}^{\text{out}} := \sum_{k \in \mathcal{W}_i' \cap \mathcal{W}_j'} p_{jk}^{\text{out}} \,, \tag{5.12}$$

where $\mathcal{V}_i'$ ($\mathcal{W}_i'$) represents the set of supplying (consuming) industries of node v_i. With the definitions in equation (5.12) we investigate if industry pairs that have established trade links with identical common supplying (consuming) industries are more likely to be positively correlated in their respective VAG. This definition is motivated by the idea that an economic trend or shock in a common neighborhood would transmit similarly to the industries. In Figure 5.7C,E it is shown that the probability distributions of common neighborhood input and output dependency for linked industry pairs in the WEPN exhibit a stronger upper tail for values of $c_{ij}^{\text{in}} > 0.2$ ($c_{ij}^{\text{out}} > 0.2$) as compared to unlinked industry pairs. The corresponding bar charts of the higher-order variables are illustrated in Figure 5.7B,D,F. The statistical differences between the distributions for linked and unlinked industries in the WEPN are summarized in Table 5.2.

In contrast to the previously discussed first-order relationships, the distributions of the common neighborhood dependency variables still show differences when we distinguish between the subsets of national and international pairs. We observe that the probabilities to obtain large values of $c_{ij}^{\bullet}$ are higher for linked than for not linked international pairs. This difference, however, is less pronounced than in the full set and in the national set of industry pairs, as can be seen in Figure 5.8 and Table 5.2

Further evidence for the particular importance of the input dependency p^{in} as well as the common neighborhood input dependency c^{in} is provided in Figure 5.9. We define the quantile range $QN = (q_1, q_2)$ as the interval of all values x of a variable such that

$$q_1 < \int_{-\infty}^{x} f(x')dx' < q_2 \tag{5.13}$$

with the probability distribution f of this variable. By utilizing the quantile ranges $QN_1 := (0.5, 0.7)$, $QN_2 := (0.7, 0.9)$ and $QN_3 := (0.9, 1)$, we observe that the distribution for industry pairs with $\text{Corr}_{ij} > 0.46$ ($\text{Corr}_{ij} > 0.2$) shows larger values for pairs in the QN_3 quantile range of the p^{in} (c^{in}) distributions as compared to pairs belonging to lower quantiles. The same result holds for the respective output variables p^{out} and c^{out} for values of $\text{Corr}_{ij} > 0.49$ ($\text{Corr}_{ij} > 0.28$).

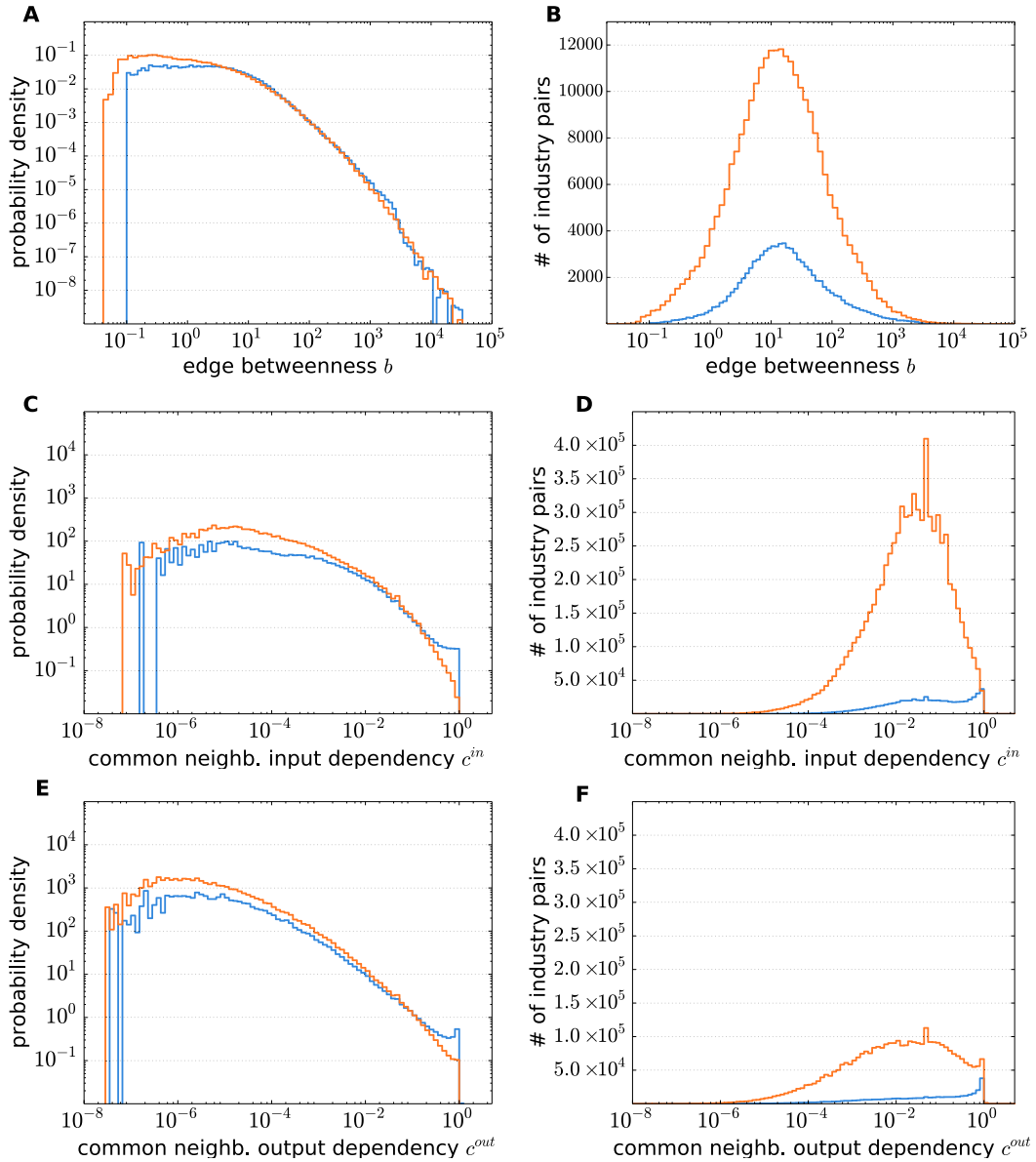


Figure 5.7.: Probability distributions of mean edge betweenness b_{ij} (A), common neighborhood input dependency c_{ij}^{in} (C) and the common neighborhood output dependency c_{ij}^{out} (E) between all pairs of industry sectors with significantly positively correlated (blue) and not significantly positively correlated (orange) VAG time series at a 0.05 significance level. The corresponding bar charts of the respective variable are shown in (B), (D) and (F).

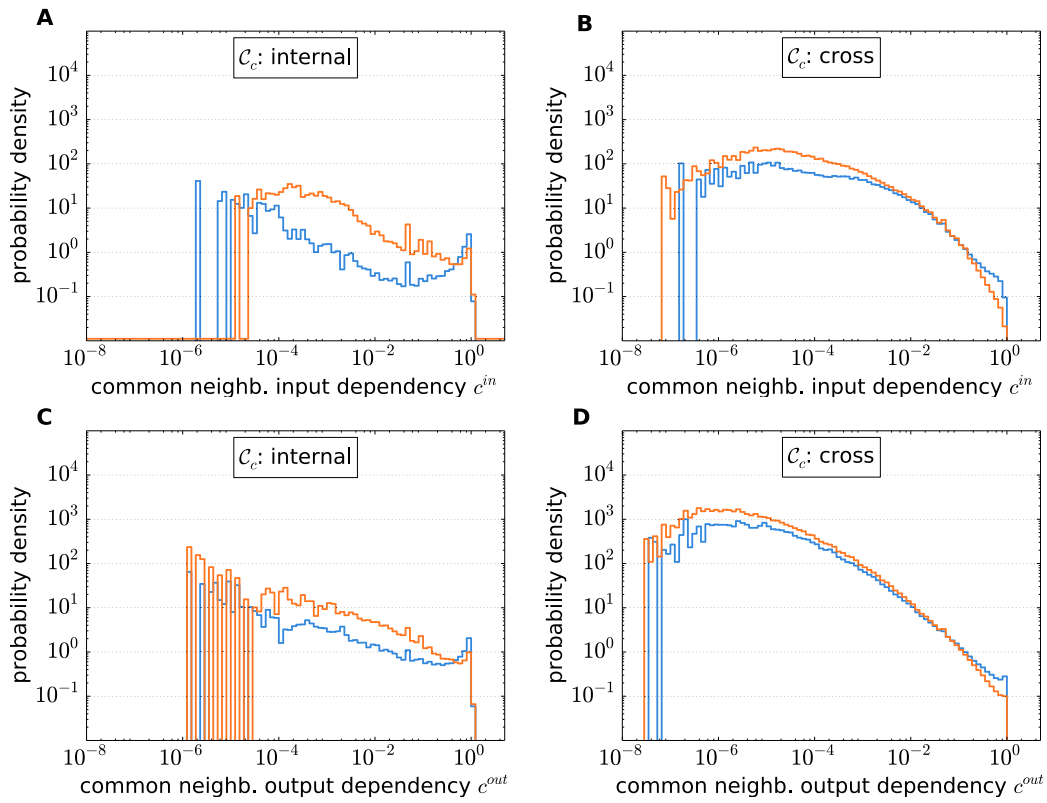


Figure 5.8.: Probability distributions of the common neighborhood input dependency c_{ij}^{in} (A,B) and the common neighborhood output dependency c_{ij}^{out} (C,D) distinguished between significantly positively correlated pairs (blue) and not significantly positively correlated pairs (orange) at a significance level of $\alpha = 0.05$. The left column depicts the distributions for internal subnetwork pairs in the national partition $\mathcal{C}_c$. The distributions of only international pairs are presented in the right column.

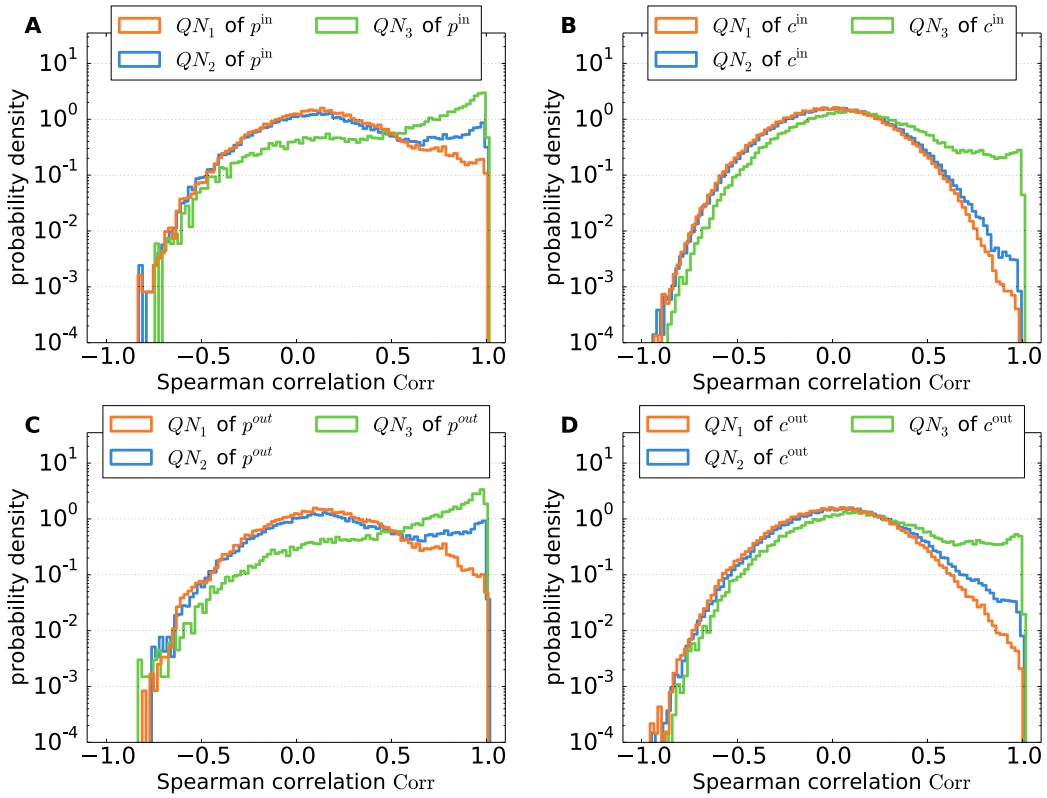


Figure 5.9.: (A,B): Probability density of the Spearman correlation between the VAGs of industry pairs in different quantile ranges $QN_\bullet$ of the input dependency p^{in}(A) and common neighborhood input dependency c^{in}(B). (C,D): The same for the output variables p^{out}(C) and c^{out}(D).

In order to further evaluate the quality of the common neighborhood dependency as an indicator for a link in the WEPN, we examine the correlation between c_{ij} and p_{ij}. First, we observe that for 15.4 % of industry pairs with significantly positively correlated VAG, the common neighborhood dependency is $c_{ij}^{\text{in}} = 0$ (for 47.9 % of the pairs it is $c_{ij}^{\text{out}} = 0$). This implies that 84.6 % of the industry pairs share at least one common supplying industry, whereas only 52.1 % share at least one common consumer, indicating a tendency towards a tree-like structure in the ITN. Both numbers are larger as compared to the 12.7 % of pairs where a direct trade relation between industries has been established. The common neighborhood input (output) dependency distribution for industry pairs with $p_{ij}^{\text{in}} = 0$ ($p_{ij}^{\text{out}} = 0$) is illustrated in Figure 5.10A (Figure 5.10C). We observe that the excess probability in the right tail for industry pairs with significant positive correlation in VAG is smaller in Figure 5.10A as in Figure 5.7C. For linked industry pairs in the WEPN with $p_{ij}^{\text{in}} > 0$ ($p_{ij}^{\text{out}} > 0$), the interrelation between the first-order and second-order dependencies is presented in Figure 5.10B (Figure 5.10D). We find that there is a considerable amount of linked industry pairs in the WEPN with high c_{ij} and low p_{ij}, i.e. 26.2 % (18.0 %) of pairs with $c_{ij}^{\text{in}} > 0.5$ and $p_{ij}^{\text{in}} < 0.001$ ($c_{ij}^{\text{out}} > 0.5$ and $p_{ij}^{\text{out}} < 0.001$). This emphasizes the importance of common neighborhoods in both supplying and consuming industries as indicators for positive correlations in VAG.

5.4.2. Network from negative correlations in VAG

Negative correlations can be expected if two industrial sectors show a similar specialization in production and therefore compete for the same suppliers or consumers. As discussed in section 3.4.2, the volumes of national trade are in general larger than the volumes of international trade. Furthermore, we have seen that the link density for the internal links in the national partition is comparatively low in the network from negative correlations in VAG (see Figure 5.4). Thus, we expect negative correlations mainly between industry sectors of different countries.

To investigate if competition between industries can be an indicator for a negative correlation in VAG, we preselect industry pairs for which the ITN topology suggests that competition between the two nodes can be likely. For this purpose, we quantify to what extent two industries v_i and v_j rely on the same international suppliers and consumers and define the *common international neighborhood* $\mathcal{VIN}^{in}$ ($\mathcal{VIN}^{out}$) as

$$\mathcal{VIN}_{ij}^{\bullet} := \mathcal{V}'_i \cap \mathcal{V}'_j \setminus \left[(\mathcal{V}'_i \cap \mathcal{V}'_{IJ}) \cup (\mathcal{V}'_j \cap \mathcal{V}'_{IJ}) \right] \tag{5.14}$$

which is illustrated in Figure 5.11. In equation (5.14), $\mathcal{V}'_i$ and $\mathcal{V}'_j$ describe the neighborhoods of nodes v_i and v_j, respectively. The neighborhoods are defined separately by either from all outgoing links and all incoming links, depending on $\bullet \in \{out, in\}$. $\mathcal{V}'_{IJ}$ describes the set of all industry nodes that are in the same country as the country of either node v_i or v_j.

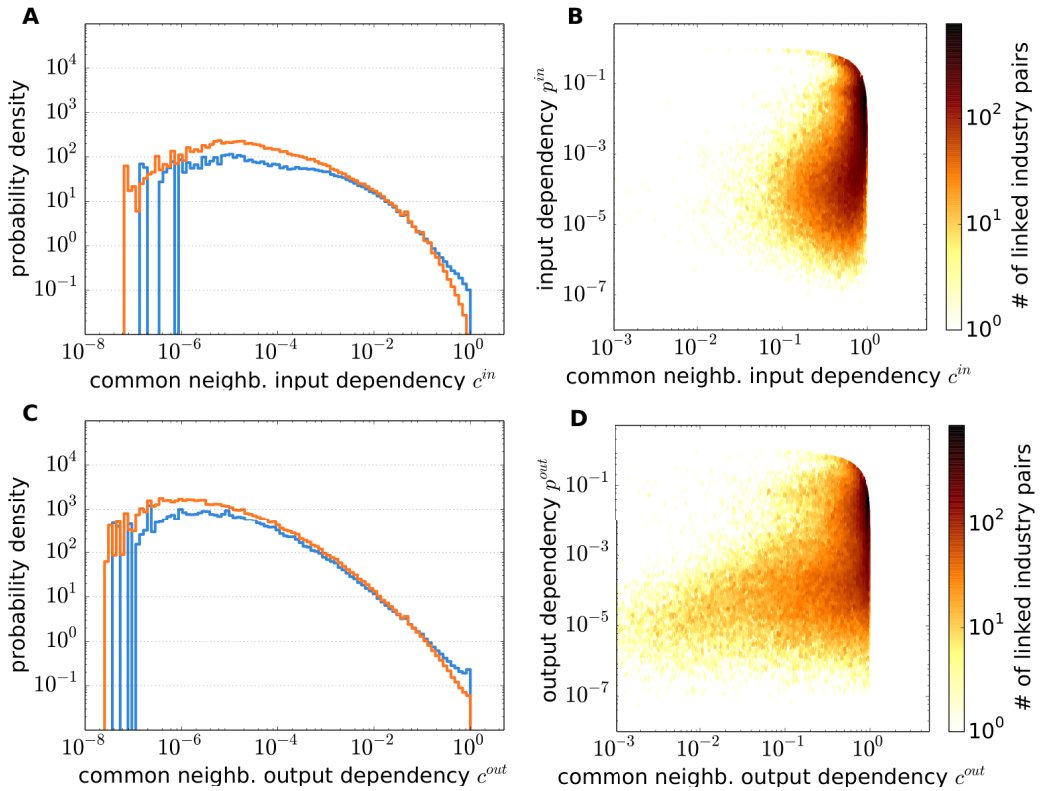


Figure 5.10.: Frequency distributions of (A,B) common neighborhood input and (C,D) output dependency for industry pairs with input (output) dependency (A) $p_{ij}^{\text{in}} = 0$, (B) $p_{ij}^{\text{in}} > 0$, (C) $p_{ij}^{\text{out}} = 0$ and (D) $p_{ij}^{\text{out}} > 0$. In panels (A,C) the blue line (orange line) shows the distribution for positively correlated (not positively correlated) pairs of sectors in VAG. In panels (B,D) only linked industry pairs in the WEPN are considered.

The common international neighborhood dependency measures are then defined as

$$cn_{ij}^{\text{in}} := \sum_{k \in \mathcal{VIN}_{ij}^{in}} p_{kj}^{\text{in}} \quad \text{and} \quad cn_{ji}^{\text{out}} := \sum_{k \in \mathcal{VIN}_{ij}^{out}} p_{jk}^{\text{out}} , \tag{5.15}$$

for the input and output, respectively.

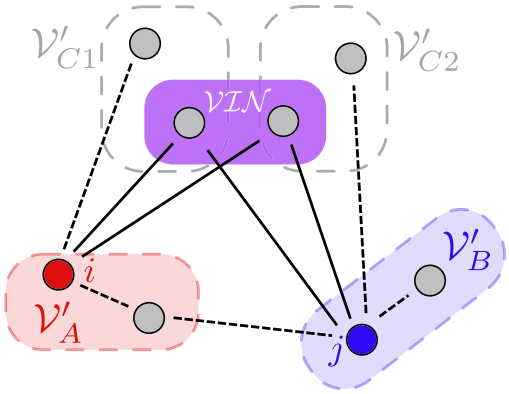


Figure 5.11.: Illustration of the common international neighborhood $\mathcal{VIN}$ (violet) of the nodes v_i and v_j that are in country $\mathcal{V}'_A$ and $\mathcal{V}'_B$, respectively. By definition, the set of nodes $\mathcal{VIN}$ can exhibit nodes from different countries $\mathcal{V}'_{C1}$ and $\mathcal{V}'_{C2}$, but not from $\mathcal{V}'_A$ and $\mathcal{V}'_B$.

When nodes in the common international neighborhood of v_i and v_j do most of their trade with either of these two industries, this suggests that v_i could benefit from a decline in v_j or vice versa. The mechanism is motivated by the assumption that the industries in $\mathcal{VIN}$ would then increase the trade with the partner which is not affected by the decline. If both v_i and v_j do most of their trade with $\mathcal{VIN}$, this effect could be particularly present. Formally, if $\mathcal{VIN}_{ij}^{out}$ is highly input dependent on the industries v_i and v_j and the two industries are at the same time highly output dependent of $\mathcal{VIN}_{ij}^{out}$, this suggests that the two nodes could be in direct competition for the identical consumers in $\mathcal{VIN}_{ij}^{out}$. Thus, we could expect negative correlations in VAG between two industries v_i and v_j when

$$c_{\mathcal{VIN},ij}^{in} := \sum_{\substack{k \in \mathcal{VIN}_{ij}^{out} \\ l \in \mathcal{V}}} \frac{w_{ik} + w_{jk}}{w_{lk}} \tag{5.16}$$

and the common international neighborhood output dependencies cn_{ij}^{out} and cn_{ji}^{out} are high. Analogously, companies compete for the same suppliers if

$$c_{\mathcal{VIN},ij}^{out} := \sum_{\substack{k \in \mathcal{VIN}_{ij}^{in} \\ l \in \mathcal{V}}} \frac{w_{ki} + w_{kj}}{w_{kl}} \tag{5.17}$$

and cn_{ij}^{in} and cn_{ji}^{in} are high.

To preselect industry pairs (i, j) for which negative correlations VAG correlations can be expected from the arguments described above, we investigate the distributions of $c_{\mathcal{VIN},ij}^{in}$ and $c_{\mathcal{VIN},ij}^{out}$ of the ITN and select all pairs that lie above the 99.5 % quantile (see Table 5.3) of the respective distribution. We preselect the pairs separately for the ITN of each year and consider then all pairs (i, j) that have been at least selected once in the ITNs between 1990 and 2011. From the subset of these considered pairs, we investigate if highly competing pairs (i, j) with a high dependency on $\mathcal{VIN}_{ij}$ are likely to be negatively correlated in VAG. Figure 5.12A (Figure 5.12B) shows the distributions of cn_{ij}^{in} (cn_{ij}^{out}) for significantly negatively correlated and not significantly negatively correlated industry pairs of the preselected subset, respectively. We observe that a high common international dependency measure in the ITN is not a good indicator for a negative correlation in VAG between the industry sectors. This holds for both the industry pairs that compete for the suppliers (cn_{ij}^{in}, Figure 5.12A) and for the pairs that compete for the consumers (cn_{ij}^{out}, Figure 5.12B).

Table 5.3.: 99.5 % quantiles of the distributions of the dependency measures of the common international neighborhood $\mathcal{VIN}$ on the nodes v_i and v_j.

variable	99.5 **% quantile**
$c_{\mathcal{VIN},ij}^{in}$	0.018
$c_{\mathcal{VIN},ij}^{out}$	0.009
$c_{\mathcal{VIN},q,ij}^{in}$	0.82
$c_{\mathcal{VIN},q,ij}^{out}$	0.54

Until now, we have not taken into account that different industrial sectors are often specialized in the production of specific goods. However, if we assume perfect complementarity, industries can substitute their required inputs exclusively from other nodes of the same sector as the sector of the original supplier. With this assumption, we further restrict the analysis to pairs (i, j) in which the nodes v_i and v_j are of the same sector q. With $\mathcal{V}'_q$ denoting then the subset of all nodes of sector q, we define the dependency of $\mathcal{VIN}_{ij}^{\bullet}$ on the two nodes on input (output) from sector q as:

$$c_{\mathcal{VIN},q,ij}^{in} := \sum_{\substack{k\in\mathcal{VIN}_{ij}^{out}\\ l\in\mathcal{V}'_q}} \frac{w_{ik}+w_{jk}}{w_{lk}} \quad \text{and} \quad c_{\mathcal{VIN},q,ij}^{out} := \sum_{\substack{k\in\mathcal{VIN}_{ij}^{in}\\ l\in\mathcal{V}'_q}} \frac{w_{ki}+w_{kj}}{w_{kl}} \,. \tag{5.18}$$

We then proceed as described above and preselect the pairs (i, j) above the 99.5 % quantile of the distributions of $c_{\mathcal{VIN},q,ij}^{in}$ in $c_{\mathcal{VIN},q,ij}^{out}$, respectively (cf. Table 5.3). Then, we investigate if the dependency measures cn_{ij}^{out} and cn_{ij}^{in} can serve as an indicator for negative correlations in VAG of the preselected pairs. From Figures 5.12C and D we observe that contrary to our suggestion, a high value of the dependency measures

does not imply a higher chance of a negative correlation in VAG between the nodes v_i and v_j.

We find that although negative correlations in VAG are observed at a higher rate than expected from the false positive rate, our proposed measures that are motivated by a competition effect in the ITN are not able to explain the occurrence of the negative correlations in general. Thus, statistically, the effect that industries in the same neighborhood follow a common trend dominates the competition effect at this level of sectoral aggregation. A problem in the detection of competition is the low statistics in the analysis with the assumption of perfect complementarity, as each sector is only compared with 185 other nodes from the same industry of different countries. Furthermore, the level of aggregation to 26 industry sectors might be too coarse such that the competition between two businesses that compete for the same consumers cannot be resolved.

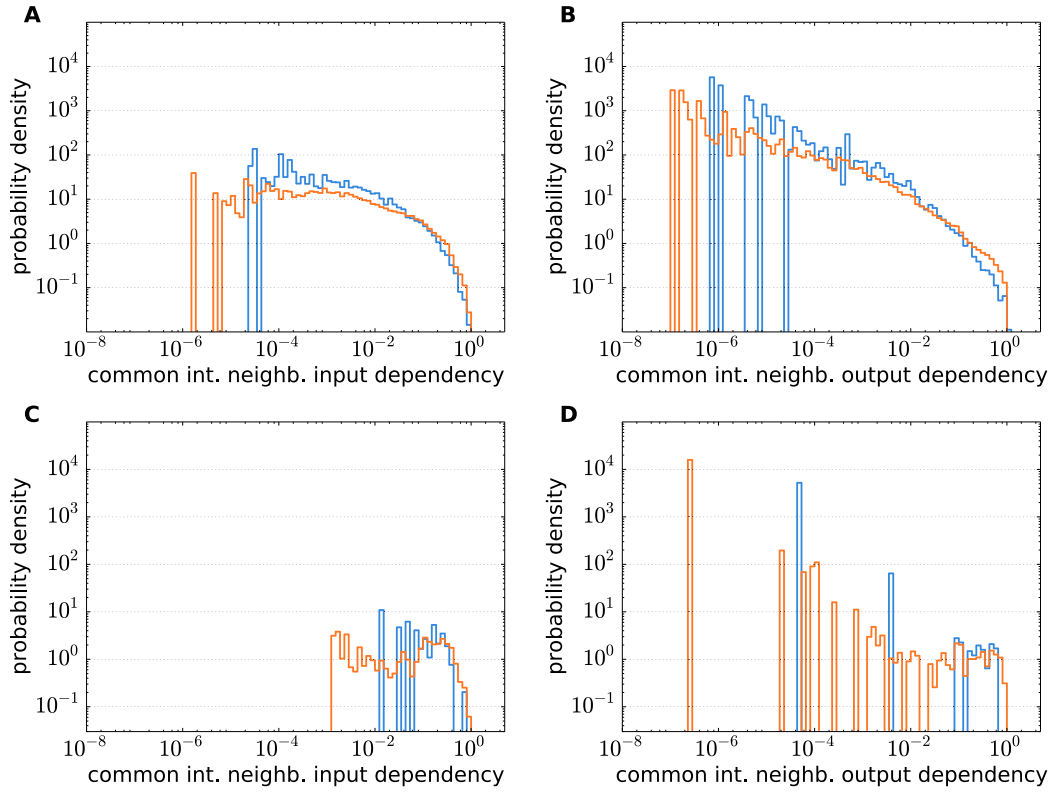


Figure 5.12.: Probability distributions of the common international neighborhood input dependency cn_{ij}^{in} (A,C) and output dependency cn_{ij}^{out} (B,D). The distributions are distinguished between significantly negatively correlated pairs (blue) and not significantly negatively correlated pairs (orange). In the top row (A,B) only pairs above the 99.5 % quantile of $c_{\mathcal{VIN},ij}^{in}$ are considered, while in the bottom row (C,D) only pairs above the 99.5 % quantile of $c_{\mathcal{VIN},q,ij}^{in}$ are considered, assuming perfect complementarity.

5.5. Discussion

We have investigated the quality of several topological properties of the ITN as indicators for both significantly positive and negative correlations in the value added growth (VAG) in the production of industries. Furthermore, we have estimated the correlation in VAG between pairs of sectors from the data by assessing the significance of Spearman's rank correlation coefficient with surrogate data. After accounting for global economic trends and crises in the data, we construct the World Economic Performance Network (WEPN). We find that although the VAG of industries from the same country are likely to be positively correlated, the community analysis reveals a more complex structure of the WEPN. Consistently, in the network from negatively correlated industry pairs we observe a low link density within the national subnetworks. On the full network there are a more significantly negatively correlated industry pairs than expected from the false positive rate.

Our results indicate that in national trade, direct trade relations between industries are a good indicator for positive correlations in the growth rate of the value added in the production of two industries. In addition to the trade volume, an even better indicator for a positive VAG correlation is obtained by taking the first-order dependencies between individual industries into account. We have demonstrated that industries with a high output dependency (i.e. little substitution opportunities in demand) are most likely to show a positive correlation in VAG with the consuming industries. Thereby, we provide evidence that changes in demand in intermediate trade have a larger potential of contagion in the ITN than changes to supplying industries.

In international trade relations, however, the first-order variables lose their predictive power towards a correlation in VAG. Thus, there is strong evidence for the importance of higher-order dependencies in the global trade network. Furthermore, 87.2 % of all industry pairs that are linked in the WEPN have not established a significant direct trade relation. From the investigated variables that take higher-order relations in the trade network into account, we find that two industries which draw their inputs from a common neighborhood are likely to have positively correlated VAG. A common neighborhood can also indicate that two nodes compete for identical consumers. However, on the subsets of preselected industry pairs with the strongest indication in the ITN topology towards being competitive industries, no general tendency towards a negative correlation can be observed. Thus, we find evidence that the effect which describes that industries in the same neighborhood follow the same trend dominates a possible competition effect at the utilized level of sectoral resolution.

The lack of observations of the competition effect in the ITN emphasizes the limitations caused by the coarse resolution at industry level. In the following chapter, we continue with the analysis of the topology of trade on the level of business firms.

Chapter 6.

Trade on the national scale: Motifs in the Japanese Business Firm Network

The industry sectors on the macroeconomic scale of input-output tables are in fact an aggregation of all businesses activities of the individual firms in the economy. It is therefore of interest, how the network topology of trade is constituted on the microscopic level of business firms. In contrast to the macroeconomic scale, microeconomic data about the business relations between different firms in the economy is not publicly available for the global economy as a whole. Japan's credit rating agencies, however, generally obtain the information about the relations between the companies that are registered in Japan. Japan is a member of the G7 group of the world's dominating economies and therefore offers one example of a leading industrialized and developed economy.

We have observed in chapter 3.4.2 how different industries show specific characteristics in the clustering coefficient on the macroeconomic scale. Here, we set these results in the microeconomic context and investigate how 3-node motifs are distributed in the *Japanese Business Firm Network (JBFN)*. We analyze methodologically the formation of 3-node motifs on a scale-free topology by employing a conceptual model. From the economic perspective, we learn about the firms' preferences in choosing their partners. In particular, we investigate how newly founded firms connect to the other members in the network and which firms in the market are likely to merge. The sections in this chapter follow in parts closely the presentation in P3.

6.1. Introduction

As we have seen, trade networks provide examples of directed networks that allow for detailed investigations on a node's functional role on the aggregated level of industries. On the microeconomic level of business networks, individual firms are interpreted as nodes [133]. In both business and trade networks nodes take specific roles within the supply chain as service or commodity providers and/or consumers that give them characteristic positions within specific trade patterns.

Motifs are understood as essential building blocks of networks that allow for a detailed characterization and classification of the system under study in many disciplines that range from biology [134, 135] to socio-economic systems [136]. Previous work has provided some first conceptual understanding of the mechanisms behind

the formation of motifs in networks [137]. It has been shown that the connection patterns of nodes with high degrees have a particularly strong influence on the motif distribution [138]. From a dynamical perspective, motifs emerge in order to optimally exploit local stability characteristics of the network [139]. However, still little is known about the dynamical origins of the motif distributions in evolving complex networks with scale-free topology [140]. The aim of this chapter is twofold: From a conceptual perspective, we strive to better understand the formation of motifs in a scale-free network and to identify and test candidate mechanisms that could generically affect the formation process of motifs in an evolving network model. On the other hand, we want to learn in the context of firm networks about the preferences of businesses in choosing their business partners from a statistical perspective.

The business network of the Japanese economy shares some common topological features of many complex networks [141, 142]. These include the scale-free degree distribution that has been empirically observed in complex networks among a variety of different disciplines, including social networks [143–146], economics [147], biology [148] to information technology [149, 150]. The emergence of the scale-free property is often explained by preferential attachment [151, 152]. The comparatively small power-law exponent in the JBFN is well modeled by a merging mechanism between firms that exhibits features that are statistically similar to processes in other sub-disciplines of physics, such as the coagulation process in aerosols [34]. Previous empirical work has observed a strongly non-trivial motif distribution among Japanese business relations [153].

As the JBFN exhibits some universal features of complex networks, it provides an interesting case for obtaining new conceptual knowledge about the formation of motifs in complex networks. Here, we study possible candidate mechanisms that reproduce the empirical distribution of 3-node motifs in the JBFN within an evolving network model. These mechanisms should leave the statistics of simpler network characteristics, such as the total number of nodes, link density and degree distribution as invariant as possible.

We start our investigations by comparing some relevant topological characteristics of a previously suggested evolving network model [34] with those of the real-world business network in Japan. The utilized model is based on three main processes: node annihilation, node creation and merging of nodes. These processes are interpreted as bankruptcies, creations of newcomer firms, and merging processes between two firms, respectively. Firstly, we identify the main drawbacks in the description of motifs in the model. In a second step, we investigate the appearances of roles and motifs in different industries of the JBFN and introduce an industry structure to the model. To influence the motif distribution in the model, we then evaluate the impact of different rules concerning the linking preferences of newcomer nodes. Furthermore, we monitor the appearance of motifs when modifying the rules for the merging process and when adding relinking possibilities to the model.

This chapter is organized as follows: We describe the dataset (section 6.2.1), the concepts to characterize the motif distributions (section 6.2.2) and the definition of the industry subnetworks (section 6.2.3). Subsequently, section 6.3 presents the

existing basic evolving network model together with three modifications to improve the representation of the empirical motif distributions of the JBFN. We then evaluate in section 6.4 the motif distributions in the real-world data and the original network model (section 6.4.1) and focus on distinctive topological characteristics within industry subnetworks (section 6.4.2) that are used to introduce subnetwork specific rules in the network model (section 6.4.3). In section 6.4.4 we assess the impact of business preferences in choosing their trade partners which can be expressed as global rules on the motif appearances. Section 6.5 summarizes our main findings. Parts of this chapter follow closely the presentation in P3.

6.2. Japanese Business Firm Network (JBFN) construction and methods

6.2.1. Dataset description

The available data for this study encompasses information about business firm relationships that have been recorded by Teikoku Databank, one of Japan's largest corporate research providers. The database comprises pairs of firms that have reported an inter-firm business relationship in the year 2009. Next to a firm's industry attribution and its date of creation, the data contains information about the reporting date of an inter-firm relationship.

We define the real-world JBFN $\mathcal{G}_r = (\mathcal{V}_r, \mathcal{E}_r)$ by drawing a link $(i, j) \in \mathcal{E}_r$ between two companies $v_i, v_j \in \mathcal{V}_r$, if commodities or services have been exchanged in return for money. Here, the direction of the link is defined to coincide with the direction of the monetary flow. As there is no quantitative information on the monetary value of the flow, we will exclusively consider such binary networks in this chapter. The resulting real-world network consists of $|\mathcal{V}_r| = 446,108$ companies with $|\mathcal{E}_r| = 2,471,689$ links. The results of this study have been cross-checked for consistency with the available data for the firm networks from 2010 to 2015. As the main results do not change significantly with the years, we restrict our following discussions to the results for the network of 2009.

6.2.2. Motif and role comparisons

In directed networks, 3-node motifs can be distinguished between 13 connected motif patterns (denoted with index μ). An overview of 3-node motif patterns and the corresponding roles of the individual nodes is presented in Figure 6.1. Within a specific motif, each node occupies one of 30 characteristic functional roles (index R). Note that, if one node is a member of various motifs in the network, this node can simultaneously take different roles.

In business networks, a node's role is related to a firm's function in the supply chain of one final product. For example, the pattern of motif $\mu = 3$ arises from a supply chain for which the firm with role $R = 6$ serves as an intermediate component manufacturer for some end product. Note that the reconstruction of the supply

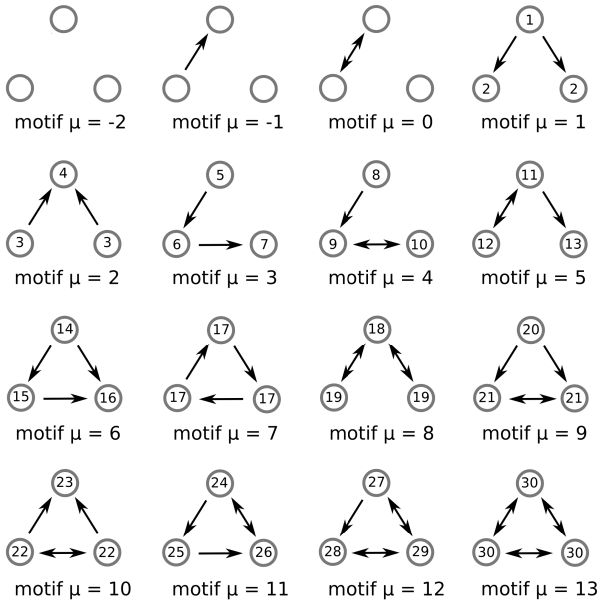


Figure 6.1.: Overview of the motif patterns μ and roles R (encircled numbers) for 3-node motifs in directed networks.

chains of individual products would still require a finer resolution of the data, as one business firm is in general active in the production of various different goods and services. Nevertheless, the role of a firm can provide meaningful interpretations. For example, if a firm specializes in the production of intermediate components, it is likely that role $R = 6$ is attributed to that firm in many motifs.

In previous studies [153], non-trivial distributions of connected motifs have been observed in the JBFN. The results have been obtained by comparing the appearances of motifs in the real-world network with expected appearances in randomized networks. These surrogates were generated by switching links such that each node keeps its number of incoming and outgoing links fixed [137]. This is a useful approach to characterize the motif appearance of real-world networks with respect to a randomized network with identical degree sequence as the original model.

In contrast, in this study we are interested in the total differences between the motif distributions in two networks $\mathcal{G}_x$ and $\mathcal{G}_y$. These differences arise partially, but not exclusively from differences in the local in- and out-degrees. Thus, the randomized null models of $\mathcal{G}_x$ and $\mathcal{G}_y$ differ in the degree distributions. Consequently, we directly compare the absolute number of motif appearances rather than their normalized deviations from the different randomized null models. This procedure avoids the problem that differences in the topological characteristics of the surrogates of the two networks will not be tracked, if the networks $\mathcal{G}_x$ and $\mathcal{G}_y$ are compared to their respective surrogates. To allow for a comparison of the absolute number of motif

appearances, we ensure that the number of nodes in the ensemble average of the model networks complies with the number of nodes in $\mathcal{G}_r$.

With x_μ denoting the number of appearances of motif pattern μ in a network $\mathcal{G}_x$ and y_μ the respective number in another network $\mathcal{G}_y$, we define the *penalty value* $\Delta P_\mu(\mathcal{G}_x, \mathcal{G}_y)$ as

$$\Delta P_\mu(\mathcal{G}_x, \mathcal{G}_y) := \frac{x_\mu - y_\mu}{\min(x_\mu, y_\mu)} \,. \tag{6.1}$$

The difference between the appearances of motif μ is normalized with respect to the network with the smaller number of appearances. This allows for a meaningful comparison between different motif patterns with different magnitudes in appearance. Equation (6.1) is anti-symmetric with respect to the exchange of networks, following $\Delta P_\mu(\mathcal{G}_x, \mathcal{G}_y) = -\Delta P_\mu(\mathcal{G}_y, \mathcal{G}_x)$. To obtain the penalty value for an ensemble of model realizations, we first average the motif frequencies of all ensemble members and then utilize equation (6.1).

6.2.3. Industry subnetworks

In trade networks, we have observed that the different directed clustering coefficients are non-trivially distributed among the different industrial sectors in the network (cf. section 3.4.2). To investigate how these characteristics are present in the microeconomic network of business firms, we assess how the industry structure in the JBFN provides information about the composition of the network's topology and the motif distribution. Let $\mathcal{G}'_\zeta = (\mathcal{V}'_\zeta, \mathcal{E}'_\zeta)$ denote the subnetwork that is induced by the subset of nodes $\mathcal{V}'_\zeta \subset \mathcal{V}_r$. Here, $\mathcal{V}'_\zeta$ contains all nodes of $\mathcal{G}_r$ that belong to an industry ζ.

To analyze differences in the topology of the individual subnetworks, we examine the degree distribution, the link density

$$\rho_\zeta := \frac{|\,\mathcal{E}'_\zeta\,|}{|\,\mathcal{V}'_\zeta\,|(|\,\mathcal{V}'_\zeta\,|-1)} \tag{6.2}$$

and the reciprocity r_ζ of the individual industries' induced subnetworks. Since the analyzed networks do not contain self-loops and as we are interested in both the link density and the reciprocity separately [154], we utilize the definition

$$r_\zeta = \frac{\sum_{i\neq j} a_{ij}^{(\zeta)} a_{ji}^{(\zeta)}}{\sum_{i\neq j} a_{ij}^{(\zeta)}} \,, \tag{6.3}$$

where the $a_{ij}^{(\zeta)}$ denote entries of the adjacency matrix of subnetwork $\mathcal{G}'_\zeta$. This definition describes the probability of the existence of an oppositely directed link when one link in the subnetwork is chosen at random.

Furthermore, we investigate the characteristic linking patterns of industries by counting the appearance of roles in motifs. To assess an industry's role characteristic,

we aggregate the appearance of role R in industry ζ by considering all nodes of this industry, obtaining the number of appearances in the industry

$$N_\zeta^R = \sum_{i \in \mathcal{G}'_\zeta} \nu_i^R \,. \tag{6.4}$$

The number of appearances of role R for one node v_i is denoted by ν_i^R. Here, motifs of the full network $\mathcal{G}$ that may extend over several industries are taken into account. In order to compare the role characteristic for industries of different size, we normalize N_ζ^R with respect to the full number of role appearances in this industry,

$$n_\zeta^R := \frac{N_\zeta^R}{\sum_{R_i} N_\zeta^{R_i}} \,. \tag{6.5}$$

The associated z-score, defined as

$$z_\zeta^R := \frac{n_\zeta^R - E_\zeta(n^R)}{\sqrt{\mathrm{Var}_\zeta(n^R)}} \,, \tag{6.6}$$

quantifies to which extent a role R is characteristic for industry ζ. Here, $E_\zeta(n^R)$ [$\mathrm{Var}_\zeta(n^R)$] denotes the mean value [variance] of the normalized appearances among the industries. A high value of z_ζ^R implies a relatively frequent appearance of role R in industry ζ.

To further improve the understanding of motifs with respect to different industries, we count the motifs in each subnetwork such that all three nodes contributing to the motif belong to the same industry. We define the subnetwork ratio of motif pattern μ and industry ζ as

$$\eta_\mu^\zeta := \frac{x_\mu^{\zeta\zeta\zeta}}{x_\mu} \,, \tag{6.7}$$

with x_μ being the number of appearances of motif μ in the full network. The number of motif appearances where all nodes belong to industry ζ is denoted by $x_\mu^{\zeta\zeta\zeta}$.

6.3. Evolving network model of the JBFN

6.3.1. Basic model

In previous work [34], it has been shown that some key topological features of the real-world JBFN $\mathcal{G}_r$ can be well approximated by an evolving network model $\mathcal{G}_m$ that incorporates a merging process of firms. The development of this model has been originally motivated by the observation of similar statistical characteristics in the mass distribution of aerosols [155]. In each evolution step of this model, a randomly selected node is either annihilated (with probability α), newly created (with probability β) or merged with another existing node in the network (with probability

$\gamma = 1 - \alpha - \beta$). When a new node is created one incoming and one outgoing link are attached to the new node. In order to reproduce the observed scale-free degree distribution of the real-world network, the selection criteria for both link creation and the merging process follow a preferential attachment rule [156]. Previous studies have demonstrated that the merging activity of firms in the real economy follows waves of market sentiment and depends on many factors, for example the size and the stock market performance of the involved firms [157, 158]. As a simplification, we do not consider these factors here and utilize the simple assumption of preferential attachment to reproduce the basic statistical characteristics of the real-world network.

The probability $\Pi_{j\to i}$ that node v_j is connected to (or merged with) node v_i reads

$$\Pi_{j\to i} = \frac{k_i + 1}{\sum_l (k_l + 1)} \, . \tag{6.8}$$

Thus, a node v_i with a high total degree $k_i = k_i^{in} + k_i^{out}$ is more likely to be chosen as business partner. With the choice of equation (6.8) we ensure that isolated nodes with $k_i = 0$ can still be chosen as new business partners.

As we will show in section 6.4.1, there are marked differences between the motif distributions obtained from the basic evolving network model as described above and those of the empirical JBFN. In order to resolve these discrepancies, we utilize the characteristics of the industry subnetworks in the real-world network $\mathcal{G}_r$ to deduce meaningful modifications of the model. Thus, we introduce an industry structure to the model and assess the impact of characteristic linking and merging preferences between firms on the appearance of motifs in the network. Specifically, we consider three types of modifications that are described in the following.

6.3.2. Introduction of industry structure

Motivated by empirical findings that are obtained by utilizing the methods introduced in section 6.2.3 and detailed later in section 6.4.2, we consider an additional industry structure as part of the network model. At the initiation of the network evolution, we randomly assign an industry to each node. The probability of a node to be associated with industry ζ is estimated from $\mathcal{G}_r$ as $|\,\mathcal{V}_\zeta\,|/|\,\mathcal{V}_{r'}\,|$, where $|\,\mathcal{V}_{r'}\,|$ denotes the total number of firms that belong to either of the following five main industries of the network (cf. section 6.4.2): construction, manufacturing, whole sales, transport & communication and services. We assign to newcomer nodes the industry with the largest population deficit that has arisen from previous annihilation and merging processes in comparison with the real-world benchmark network.

In the original model, a new node is created with exactly one incoming and outgoing link each. Here, we modify the connection pattern of newcomers and introduce two characteristic connection patterns for each industry of which one is randomly chosen at node creation with equal probability. This procedure represents a simplification of the patterns observed in the real-world network as will be discussed later in section 6.4.3. However, by restricting the modification to two linking patterns with equal probability, a clear understanding of the influence of linking preferences of newcomer firms on

the resulting motif distribution can be obtained. Furthermore, we reduce the risk of over-fitting the model by minimizing the introduction of additional parameters. The choice of the industry specific connection pattern is motivated by the results of the subnetwork analysis of the real-world network $\mathcal{G}_r$, for which the obtained specific connection patterns for each industry will be presented in section 6.4.3. We refer to the model that includes the industry structure as $\mathcal{G}_{ind}$ in the remainder of this study.

6.3.3. Δk-rule for the merging process

In section 6.4.1, we will report and discuss an observed excess frequency of motifs with bidirectional links in the model $\mathcal{G}_m$ as compared to $\mathcal{G}_r$. To account for this effect, we design a rule to reduce this overshoot in the model. Bidirectional links occur in the merging process when two companies merge that share the same business partner. When one of the merging firms serves as a supplier and the other as a consumer to the shared partner, a bidirectional link will be established in the merging process. Thus, an intuitive rule to reduce the appearance of motifs with bidirectional links is based on the local degree properties of each node. We denote the difference of a node's in-degree and out-degree as $\Delta k_i = k_i^{in} - k_i^{out}$. Following the above argument, we expect a reduction of bidirectional links if merging processes between firms v_i and v_j with $\Delta k_i \cdot \Delta k_j < 0$ are forbidden. This rule, which we refer to as the Δk-rule, can be interpreted such that two firms with similar shares of their input/output allocations are more likely to merge. With the production function describing a firm's output depending on its input, two firms with similar production functions will exhibit the same sign in Δk_i, allowing for a merging process. We indicate network models $\mathcal{G}_\bullet$ that include the Δk-rule with an additional subscript Δk.

6.3.4. Substitution of trading partners

Finally, we investigate a third modification of the model that allows companies to substitute their trading partners. Hence, we introduce a relinking process with probability δ in the evolution step. The probabilities of the processes during network evolution are then adjusted to satisfy $\alpha' + \beta' + \gamma' + \delta = 1$. Specifically, the probabilities of the annihilation, creation and merging processes are modified such that the relative ratios between the probabilities of the processes stay constant, i.e. $\alpha' = (1-\delta)\cdot\alpha$ and analogous for β' and γ'. During relinking, a business firm v_i and one link are chosen at random. We then consider this firm substituting its business relation in favor of a firm that shares a common business partner. Thus, the new business partner v_l of node v_i is found by following a path from node v_i to another node of length 2. With the symmetrized adjacency matrix with entries $a'_{ij} = a'_{ji} = \max(a_{ij}, a_{ji})$, the probability that the link is reconnected to node v_l is then given by

$$\Pi_{i,l} = \frac{(\sum_j a'_{ij} a'_{jl})(1 - a'_{il})}{\sum_m (\sum_j a'_{ij} a'_{jm})(1 - a'_{im})} \,. \tag{6.9}$$

Equation (6.9) accounts for the assumption that more common business relations make a reconnection to node v_l more likely. Models $\mathcal{G}_\bullet$ that include the relinking process are denoted with the additional subscript l.

6.4. Topology and motifs in the JBFN

6.4.1. Motif appearances in the real-world network and basic model

Previous work has shown that the model $\mathcal{G}_m$ with a choice of probabilities $\alpha = 0.2$, $\beta = 0.5$ and $\gamma = 0.3$ (for the processes of annihilation, creation and merging, respectively) approximately reproduces the scale-free degree distribution from the real-world network, following a power law $\propto k^{-\alpha}$ with an exponent of $\alpha \approx 1.4$ [133]. Despite changes in the number of firms and links, it has been concluded that some basic characteristics such as the degree distribution are robust in the JBFN from 1994 to 2014 [133]. We simulate an ensemble of 100 networks with approximately the same number of nodes $|\,\mathcal{V}_m\,|$ in the model as in $\mathcal{G}_r$. From the different realizations, we obtain a network ensemble with $|\,\mathcal{V}_m\,| = 450,000 \pm 6,000$ nodes and $|\,\mathcal{E}_m\,| = 1,090,000 \pm 20,000$ links (denoting the ensemble average and standard deviation, respectively). Note that due to the probabilistic nature of the network model, it is not possible to fix both the number of nodes and the link density at exactly the desired values corresponding to those of the real-world network.

Figure 6.2 illustrates the motif distribution for the real-world network $\mathcal{G}_r$ and the basic evolving network model $\mathcal{G}_m$. Note that many patterns of connected motifs are more frequently observed in $\mathcal{G}_m$ despite its lower link density ($\rho_m < \rho_r$). To further quantify the impact of the link density difference, we construct an Erdős-Rényi type random graph ensemble $\mathcal{G}_r^{ER}$ such that the link densities for both uni- and bidirectional links correspond to the respective densities in the real-world network. This is achieved by drawing uni- and bidirectional links with the respective probability estimated from the real-world network separately, with $\rho_{r;uni} = 1.14 \cdot 10^{-5}$ and $\rho_{r;bi} = 8.2 \cdot 10^{-7}$. $\mathcal{G}_m^{ER}$ is similarly defined for the modeled network, with $\rho_{m;uni} = 1.06 \cdot 10^{-5}$ and $\rho_{m;bi} = 9.4 \cdot 10^{-7}$. We simulate 100 realizations for the Erdős-Rényi type random graphs. As shown in Figure 6.2, the penalty values $\Delta P_\mu(\mathcal{G}_m, \mathcal{G}_r)$ are in general smaller than the corresponding values $\Delta P_\mu(\mathcal{G}_m^{ER}, \mathcal{G}_r^{ER})$ between the random networks with identical link density. This observation arises from the similarity in the degree distribution between data and model, both following a power-law $\propto k^{-\alpha}$ with almost identical exponent. For large number of nodes, the degree distributions of Erdős-Rényi type random graphs follow a Poisson distribution

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \,. \tag{6.10}$$

These distributions show different values of their characteristic parameters which are uniquely determined by the respective link density and thus by the average degree $\langle k \rangle$.

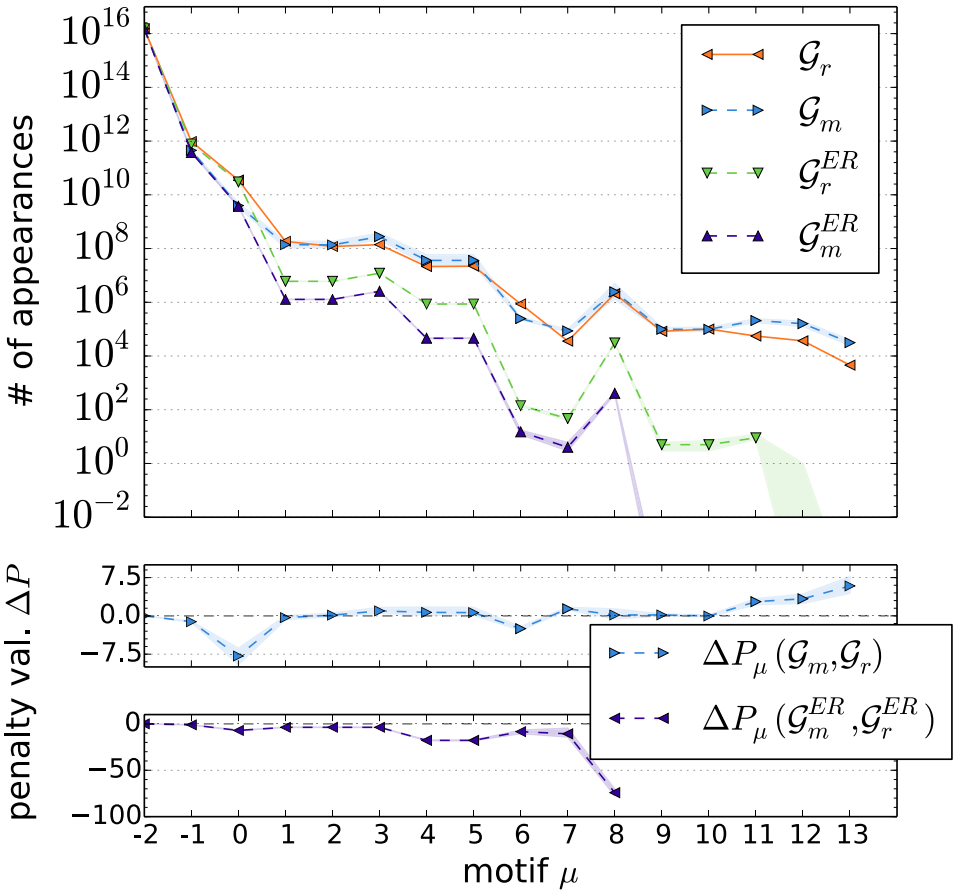


Figure 6.2.: Number of appearances of motif patterns in the real-world network $\mathcal{G}_r$ (orange solid), the original evolving network model $\mathcal{G}_m$ (blue dashed) and in random graphs $\mathcal{G}_r^{ER}$ and $\mathcal{G}_m^{ER}$ with corresponding densities of uni- and bidirectional links as observed in the data (green dashed) and in the model (violet dashed), respectively. For the models, the median values are represented as markers and the upper (lower) quartiles are illustrated by the upper (lower) shaded boundary. The median and quartiles are estimated from the respective ensemble of model realizations. The two plots at the bottom show the penalty values for $\Delta P_\mu(\mathcal{G}_m, \mathcal{G}_r)$ (blue dashed) and for the random graphs $\Delta P_\mu(\mathcal{G}_m^{ER}, \mathcal{G}_r^{ER})$ (violet dashed).

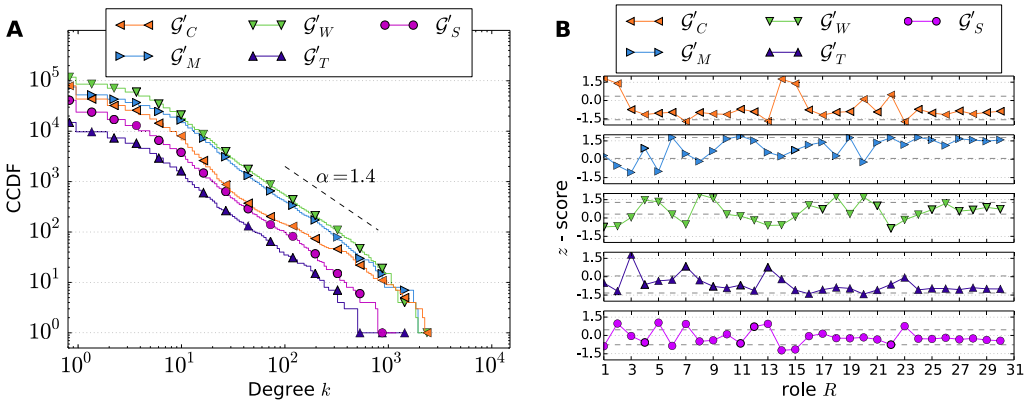


Figure 6.3.: Topological properties of the industry subnetworks in the JBFN. (A) Complementary cumulative distribution function of the total degree k_i in the induced subnetworks $\mathcal{G}'_\zeta$ for different industry categories ζ, including the sectors of C = construction, M = manufacturing, W = whole sales, T = transport & communication and S = services. (B) Role analysis of industries: z-scores as defined in equation (6.6). The values of $E(z_\zeta) \pm \sqrt{\mathrm{Var}(z_\zeta)}$ for each industry ζ are indicated by gray lines.

From Figure 6.2, we observe that for sparsely connected motif patterns, the evolving network model offers a good description of the motif distribution of the data. However, the model considerably overestimates the appearances of dense motifs with bidirectional links such as $\mu = 11$ to 13, especially $\Delta P_{13}(\mathcal{G}_r, \mathcal{G}_m) = 5.8^{+1.8}_{-1.4}$. Here, the central penalty value refers to the median of the ensemble. The upper (lower) error bands denote the differences of the upper (lower) ensemble quartiles of ΔP_μ from the corresponding median. As we will demonstrate in section 6.4.4, the large value of ΔP_{13} can be attributed to the creation of bidirectional links during the merging process in the model.

6.4.2. Industry subnetworks in the business network

In the real-world network $\mathcal{G}_r$, most firms belong to either of the five industry subnetworks with highest cardinality: whole sales ($\mathcal{G}'_W$, with $|\mathcal{V}'_W|/|\mathcal{V}_r| = 0.329$), construction ($\mathcal{G}'_C$, 0.22), manufacturing ($\mathcal{G}'_M$, 0.215), services ($\mathcal{G}'_S$, 0.145) and transport & communication ($\mathcal{G}'_T$, 0.047). We focus in this study on these main industries, as they cover 95.6 % of all nodes in the network. Figure 6.3A shows the degree distributions of the subnetworks. From visual inspection, we observe that in most industries this distribution resembles a power-law with an exponent that is similar to the exponent of $\alpha \approx 1.4$ in the full network. In the construction sector, however, the distribution does not obey a power-law due to the lack of middle-sized firms. This underlines the uniqueness in the establishment of business relations between construction firms in Japan, which is supported by findings from previous studies [159].

We further analyze the characteristic linking patterns of industries by counting the appearances of roles in motifs. The obtained z-scores are determined as described

Table 6.1.: Overview on some topological characteristics in the key industries of the Japanese economy and illustration of the introduced linking pattern in the modified model $\mathcal{G}_{ind}$. Depending on a newcomer's industry, one of the two patterns is chosen at random with equal probability at node creation.

subnetwork	$\mathbf{r}_\zeta$	ρ_ζ	**characteristics of dominant roles**	**pattern in $\mathcal{G}_{\mathbf{ind}}$** (I)	(II)
$\mathcal{G}'_C$ Construction	0.04	$1.9 \cdot 10^{-5}$	both in- and outgoing links	$\rightarrow \bullet \rightarrow$	$\rightarrow \bullet$
$\mathcal{G}'_M$ Manufacturing	0.11	$3.3 \cdot 10^{-5}$	intermediate manufacturer	$\rightarrow \bullet \rightarrow$	$\leftrightarrow \bullet \leftarrow$
$\mathcal{G}'_W$ Whole Sales	0.06	$2.0 \cdot 10^{-5}$	more out- than incoming links	$\leftarrow \bullet \rightarrow$	$\rightarrow \bullet$
$\mathcal{G}'_T$ Transport & Comm.	0.10	$9.0 \cdot 10^{-5}$	unidirectional links	$\rightarrow \bullet \leftarrow$	$\leftarrow \bullet \rightarrow$
$\mathcal{G}'_S$ Service	0.07	$2.2 \cdot 10^{-5}$	incoming links	$\rightarrow \bullet \leftarrow$	$\rightarrow \bullet$

in equation (6.6)and are presented in Figure 6.3B. We observe that manufacturing firms are more likely to occupy roles as a provider of intermediate goods ($R = 6$) and roles with a bidirectional link to one partner. Business firms in the service sector, however, are more prominent in roles with incoming links ($R = 2, 7$ and 13). Figure 6.4 illustrates excerpts of subnetworks that consist of representative business firms with a characteristic role for their industry. Here, the representative firm in the construction sector (Figure 6.4A) occupies role $R = 1$ by consuming goods and services from other firms. The providers are primarily firms from the construction or whole sales sector. In the example from the manufacturing industry (Figure 6.4B), the considered firm exhibits mainly outgoing links to the whole sales sector. This indicates the requirement of inputs for production delivered by that sector. The large number of incoming links from other manufacturing companies hint towards the firm's role as a manufacturer of intermediate goods. In the whole sales sector (Figure 6.4C) the selected firm with role $R = 18$ has established bidirectional relationships with other whole sales and manufacturing businesses. A different pattern is observed for the transport & communication and service sectors. Here, nodes with characteristic roles often exhibit a low degree but are connected to hubs in other industries. In Figure 6.4D, many firms from the transport & communication sector consume goods or services from the same manufacturing partner. Finally, in Figure 6.4E, the selected firm provides services to a manufacturing business that has many clients in the whole sales sector. A summary of the characteristic topological features in the industry-induced subnetworks, including the link density ρ_ζ and the reciprocity r_ζ, is presented in Table 6.1.

To investigate if specific industries exhibit characteristic motif patterns, we utilize the subnetwork ratio of industries as defined in equation (6.7). We observe from Figure 6.5 that 16.6 % of the appearances of the V-shaped motif $\mu = 2$ in the full network are restricted to the whole sales sector only. In turn, the manufacturing sector is dominant in densely connected motifs with bidirectional links. This finding indicates that firms in the manufacturing sector trade intermediate goods among each other more often than firms in other sectors.

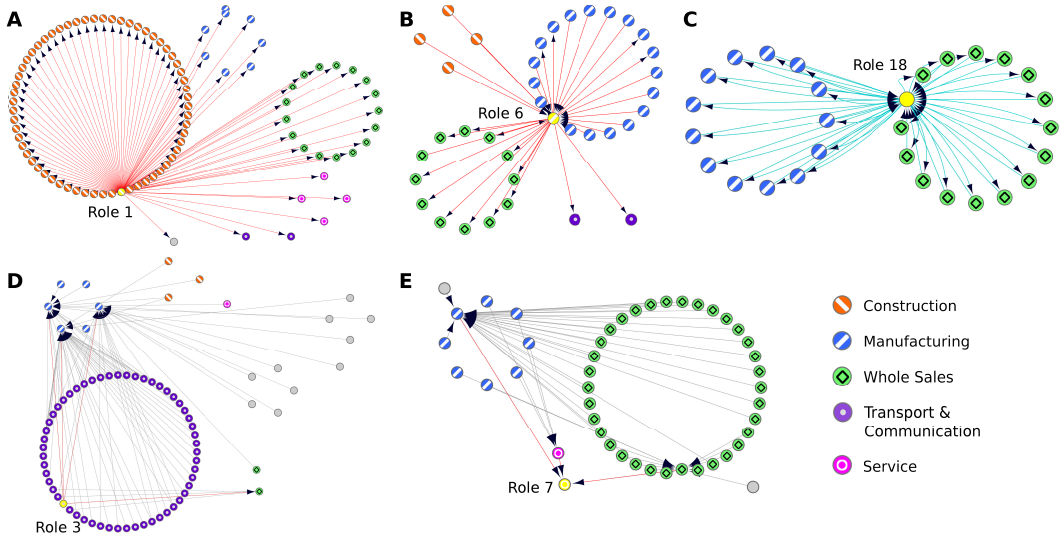


Figure 6.4.: Network excerpts around exemplary nodes with characteristic industry roles. The exemplary nodes from the construction (A), manufacturing (B), whole sales (C), transport & communication (D) and service (E) sectors are marked yellow, respectively. Unidirectional (bidirectional) links connected to the selected nodes are plotted in red (light blue). Gray edges depict unidirectional links that are not connected to the selected nodes.

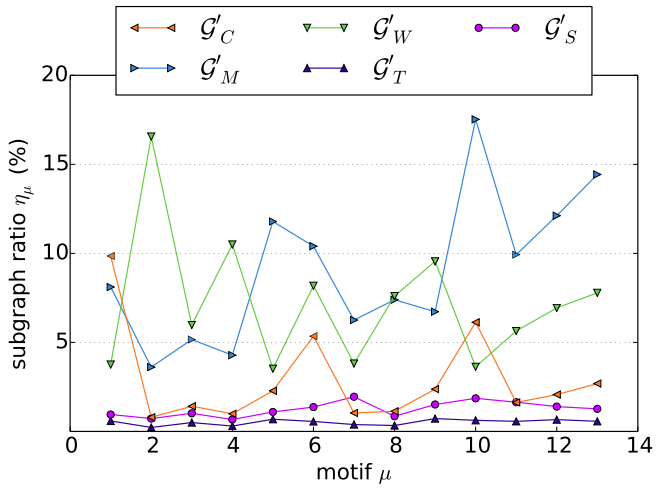


Figure 6.5.: Subnetwork ratio η_μ of motif patterns with respect to different industries' subnetworks $\mathcal{G}'_\zeta$ in the JBFN $\mathcal{G}_r$.

6.4.3. Impact of the industry structure on the motif appearances

With the results presented above, we introduce an industry structure to the model and modify the linking pattern at the creation process of a new node as described in section 6.3.2.

The considered linking patterns in $\mathcal{G}_{ind}$ and the results of the subnetwork analysis of $\mathcal{G}_r$ are summarized in Table 6.1. The comparatively high value of the reciprocity in the manufacturing sector $\mathcal{G}'_M$ indicates a higher density of bidirectional links as compared to most other industries. Therefore, we introduce the possibility of a bidirectional link for newcomer nodes from the manufacturing sector. We observe that $z^r_M > \mathrm{E}(z_M) + \sqrt{\mathrm{Var}(z_M)}$ for $R = 6, 11, 19$ and 24 (cf. Figure 6.3). Except for $R = 19$ these roles describe the production of intermediate goods. Thus, we keep the pattern of one incoming and one outgoing link for manufacturing nodes. In the transport & communication industry $\mathcal{G}'_T$, the reciprocity also shows a comparatively high value. However, we do not include a bidirectional link for $\mathcal{G}'_T$ due to the infrequent appearance of roles with these links (i.e. roles $R =$ 9-12 and $R =$ 26-30, see Figure 6.3). This illustrates the importance of the connection to high-degree firms in other industries, as shown in Figure 6.4D. As the roles $R = 3, 7, 13$ with $z^r_T > \mathrm{E}(z_T) + \sqrt{\mathrm{Var}(z_T)}$ exhibit either incoming or outgoing links, we introduce the patterns with two links of identical directions (either incoming or outgoing) to the node for $\mathcal{G}'_T$. Motivated by the lower link density ρ_ζ in $\mathcal{G}'_C$, $\mathcal{G}'_W$ and $\mathcal{G}'_S$, we introduce the pattern with a single unidirectional link only for newcomer nodes of these industries. As most motif patterns μ appear more frequently in $\mathcal{G}_m$ than in $\mathcal{G}_r$ (cf. Figure 6.2), we do not consider the possibility of adding more than two links to newcomers. In the whole sales sector $\mathcal{G}'_W$, we observe more outgoing than incoming links for the characteristic roles $R = 4, 5, 8, 9, 18$ and 20. On the other hand, we see a strong dominance of roles $R = 2, 5, 7, 12$ and 13 with incoming links in the service sector $\mathcal{G}'_S$. In $\mathcal{G}'_C$ both incoming and outgoing links are almost equally present in the roles with $z_C > \mathrm{E}(z_C) + \sqrt{\mathrm{Var}(z_C)}$. We consider these results in the directionality of links in the selected patterns as summarized in Table 6.1.

To compare the introduced connection patterns with the empirical data, we analyze the directionality distribution of newly established firms in the Japanese economy from $\mathcal{G}_r$. For this purpose, we define the fraction of incoming initial links θ^{in}_i for firm v_i as

$$\theta^{in}_i = \frac{\kappa^{in}_i}{\kappa_i} \,. \tag{6.11}$$

Here, κ_i denotes the number of initially established links of firm v_i and κ^{in}_i describes the number of incoming links among them. Motivated by a comparison of the real-world data with the original evolving network model we consider $\kappa_i = 2$. There are, however, firms in the database with more than 2 links that have been reported in the same month as the establishment of the first link. In these cases, the first two links cannot be uniquely identified and we set the value of $\kappa_i > 2$ such that all reported links of this month are considered. In cases when the first 2 links of a newcomer firm

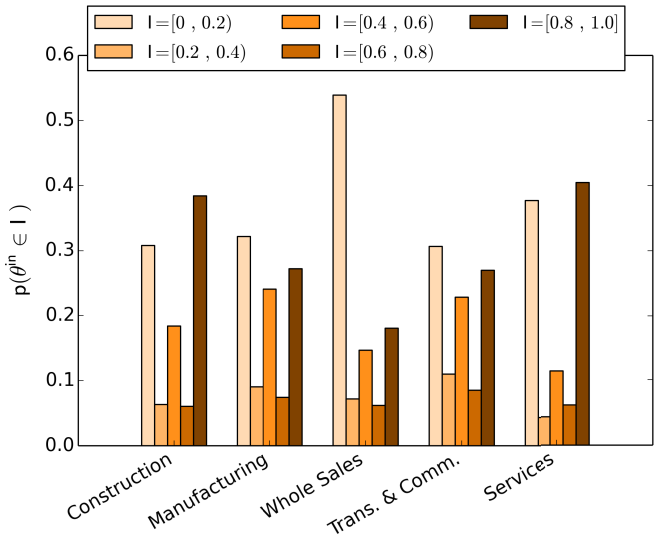


Figure 6.6.: Distribution of the fraction of incoming initial links θ^{in} for key industries in the JBFN. The bars indicate the probabilities $p(\theta^{in} \in I)$ of values in the intervals I defined as described in the legend.

can be obtained from the data, the statistic of θ^{in} in equation (6.11) allows for a direct evaluation of the linking pattern of newcomers in the real-world network that can be compared with the introduced linking pattern (I) of $\mathcal{G}_{ind}$ in Table 6.1. If $\kappa_i > 2$, the method still provides insights to the question if the first links of newcomers are predominantly incoming or outgoing.

The distribution of θ^{in} is illustrated in Figure 6.6. We observe a strong tendency of firms in the whole sales sector to firstly establish outgoing links to other companies. Firms in this industry first buy the products before further distributing them among other companies or final consumers. In the construction and service sectors, new participants in the market tend to start as providers, establishing incoming links first. Although the observed connection patterns are less distinctive than in the model, these findings are in line with the introduced simplified pattern scheme in Table 6.1.

The impact of the introduction of the industry structure to the model $\mathcal{G}'_{ind}$ as described above is evaluated in Figure 6.7. The figure shows the subnetwork ratio η^{ζ}_{μ} at the level of industry subnetworks for $\mathcal{G}_r$ and for the model $\mathcal{G}_{ind}$ that incorporates the modified industry-dependent linking patterns at node creation as described above. Some empirical characteristics of the real-world network $\mathcal{G}_r$ are reproduced by $\mathcal{G}_{ind}$. These include the frequent appearance of a central node with two incoming links (motif $\mu = 2$) in the whole sales industry and a predominance of motifs with a central node and one bidirectional link (motif $\mu = 5$) in the manufacturing sector. On the other hand, the modified linking patterns at node creation do not have a large influence on densely connected motifs with bidirectional links (motifs $\mu = 9$-13). Furthermore, the impact of the modifications in $\mathcal{G}_{ind}$ on the distribution among the

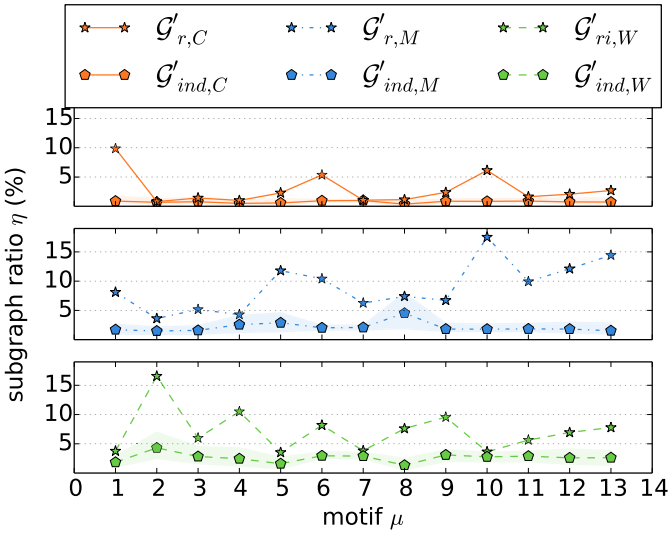


Figure 6.7.: Impact of model $\mathcal{G}_{ind}$ on the motif appearance within industry subnetworks. The figure shows the subnetwork ratio η of motif patterns for the three industry subnetworks with the highest numbers of firms in the real-world network $\mathcal{G}_r$ and in the modified model with industry structure $\mathcal{G}_{ind}$.

industries is still low as compared to the differences in the real-world network. This result underlines the importance of the merging process in influencing the motif appearance.

The impact of the introduction of the industry structure to the model on the full network is illustrated in Figure 6.8. We observe that the industry structure improves the description of motifs $\mu = 0, 3, 7, 11, 12$ and 13, whereas the absolute penalty value increases for motifs $\mu = 1, 2, 4, 5, 6$ and 8. However, the improvements outweigh the negative effects (see Table 6.2 that shows the sum of $|\Delta P| := \sum_\mu |\Delta P_\mu|$ over all 16 motifs for different model variants). Nevertheless, despite a positive impact on the densely connected motifs $\mu = 11$-13, further modifications of the model are required to more accurately describe the motif distribution of the real-world network.

6.4.4. Impacts of merging preferences and substitution of links on the motif appearances

From the findings discussed above, we deduce that a reduction of the observed excess frequency of densely connected motifs in the model as compared to $\mathcal{G}_r$ requires an additional modification of the merging rule. Specifically, this excess cannot be simply explained by the introduction of an industry structure to the model alone (see Figure 6.8). Therefore, we utilize in the following the Δk-rule (cf. section 6.3.3). Thus, only pairs of firms with similarities in their input/output allocations merge in this model variant. We apply this rule in addition to the introduction of the industry structure and investigate the results of the resulting model $\mathcal{G}_{ind,\Delta k}$.

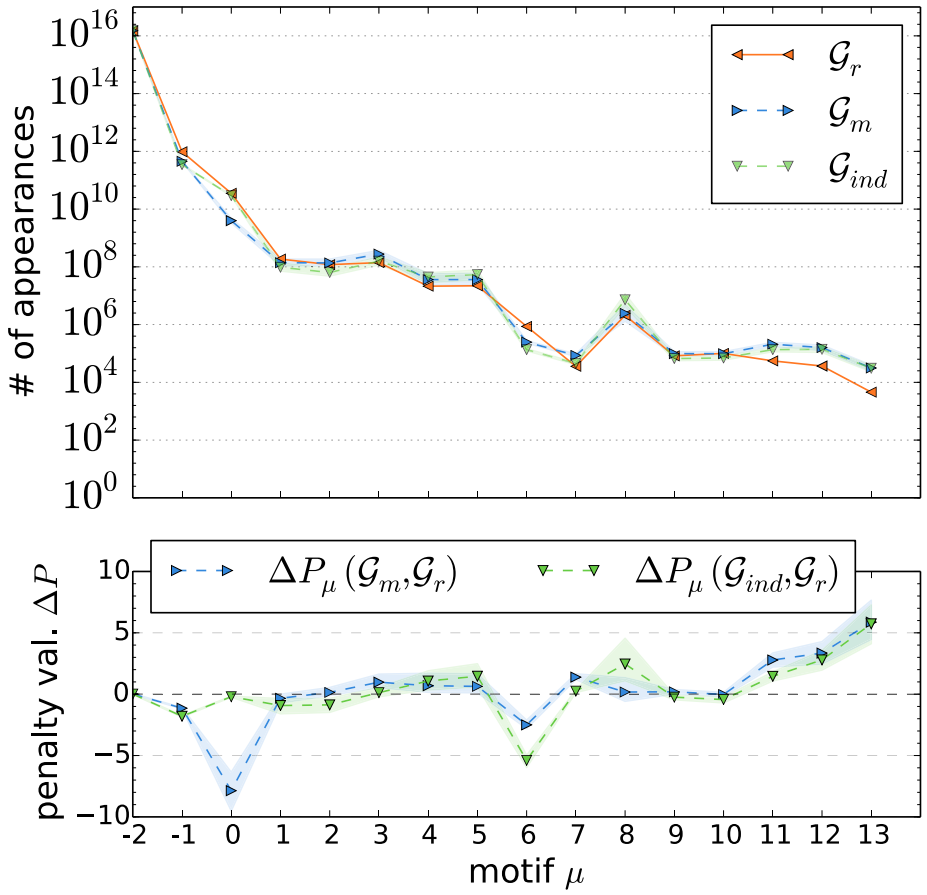


Figure 6.8.: Impact of model $\mathcal{G}_{ind}$ on the motif distributions in the full network. Top: Absolute frequencies of motif appearances in the real-world network $\mathcal{G}_r$ (solid blue), the model $\mathcal{G}_m$ (red) and the modified model with and industry structure $\mathcal{G}_{ind}$ (light blue). Bottom: Penalty values ΔP_μ as defined in equation (6.1) for two model variants with respect to the real network. For the models, the median (markers) and the respective upper and lower quartile (shaded boundaries) of the ensemble realizations are shown.

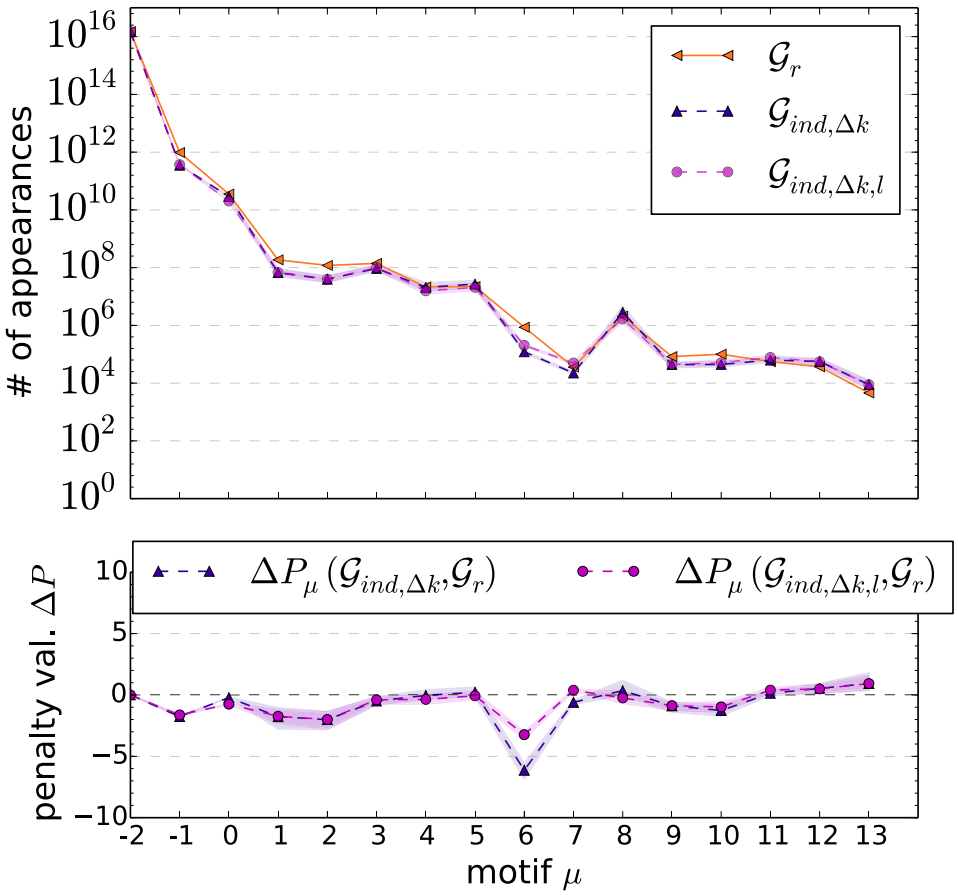


Figure 6.9.: Impact of the Δk-rule and the introduction of relinking possibilities (l) on the motif distributions in the full network. As in Figure 6.8 for the two model variants $\mathcal{G}_{ind,\Delta k}$ and $\mathcal{G}_{ind,\Delta k,l}$ that both involve the Δk-rule.

Table 6.2.: Sum of the penalty values for all motif patterns in the investigated model variants.

model	$\mathcal{G}_m$	$\mathcal{G}_{ind}$	$\mathcal{G}_{ind,\Delta k}$	$\mathcal{G}_{ind,\Delta k,l}$	$\mathcal{G}_{\Delta k}$	$\mathcal{G}_l$	$\mathcal{G}_{\Delta k,l}$	$\mathcal{G}_{ind,l}$
$\lvert\Delta P(\mathcal{G}_r, \mathcal{G}'_\bullet)\rvert$	28^{+11}_{-6}	25^{+10}_{-8}	17^{+7}_{-4}	14^{+5}_{-4}	33^{+15}_{-10}	29^{+9}_{-5}	32^{+13}_{-9}	21^{+8}_{-6}

Figure 6.9 shows the total numbers of motif appearances (top) and the penalty values ΔP_μ (bottom) for the modified model $\mathcal{G}_{ind,\Delta k}$ that includes both the industry structure and the Δk-rule. We observe that the introduction of the Δk-rule has a substantial impact on motifs with bidirectional links, greatly reducing the number of appearances of motifs $\mu = 8, 11, 12$ and 13. This leads to a much better description of the data in these motif patterns, with a penalty value of $\Delta P_{13}(\mathcal{G}_r, \mathcal{G}_{ind,\Delta k}) = 0.9^{+0.8}_{-0.5}$ in motif $\mu = 13$. This improvement of the model leads to the hypothesis that pairs of firms that allocate their input/output similarly show a higher probability to merge.

As shown in section 6.4.3 the introduction of the industry sector results in a better description, especially of unconnected motifs. This is mainly achieved by introducing the bidirectional link pattern to manufacturing firms. For the sparsely connected motifs $\mu =$ 1-5 the models $\mathcal{G}_m$ and $\mathcal{G}_{ind}$ describe the motif appearance as observed in the real-world network almost equally well (cf. Figure 6.8). In turn, this rule provides a more accurate description of unconnected motifs with bidirectional links ($\mu = 0$) and reduces the deficit of densely connected motifs $\mu =$ 11-13. However, for the fully connected motif with unidirectional links ($\mu = 6$) the introduction of the industry structure results in a less accurate modeling of the real-world network. This can be attributed to the introduction of nodes in three industries that receive only one link at creation.

Furthermore, the deficit of appearances of motif $\mu = 6$ can be explained by the lack of relinking possibilities for companies in the models $\mathcal{G}_m, \mathcal{G}_{ind}$, and $\mathcal{G}_{ind,\Delta k}$. As described in section 6.3.4 and equation (6.9), we finally investigate the impact of additional relinking possibilities on the motif distribution by considering substitution opportunities for companies. In the model variant $\mathcal{G}_{ind,\Delta k,l}$, we assume that companies substitute their relations in favor of common business partners. To find a suitable value for the relinking probability δ in model $\mathcal{G}_{ind,\Delta k,l}$, we investigated the impact on the motif distribution for different values of δ. The possibility to substitute business relations improves the description of motif pattern $\mu = 6$ in the model. A high value of δ results in a high frequency of this pattern. However, the modification l also increases the frequency of pattern $\mu = 7$ and decreases the frequency of $\mu = 0$ and 8. With these opposing effects on the motif distribution, we observe the lowest sum of the penalty values for $\delta = 0.3$. At this probability the largest effect on motif $\mu = 6$ is obtained, while leaving the impacts on other motif patterns relatively small. Compared to $\mathcal{G}_{ind,\Delta k}$, the penalty value $|\Delta P_6|$ considerably decreases without increasing $|\Delta P_\mu|$ for $\mu \neq 6$. The corresponding motif distributions for the different values of δ are illustrated in the Figures C.1 and C.2 in the appendix. With $\delta = 0.3$, the median value of $\sum_\mu |\Delta P_\mu|$ decreases from 17^{+7}_{-4} for $\mathcal{G}_{ind,\Delta k}$ to 14^{+5}_{-4} for $\mathcal{G}_{ind,\Delta k,l}$.

To compare the description of the motif distributions among all variants of the evolving network model, we calculate the sum $|\Delta P| := \sum_\mu |\Delta P_\mu|$ over the 16 motif patterns μ for all model modifications in Table 6.2. The table includes all combinations of the discussed modifications. We see that the Δk-rule and the introduction of relinking possibilities alone do not decrease the penalty value $|\Delta P|$ compared to the original model $\mathcal{G}_m$. This can be attributed to the fact that while these models improve the descriptions of motifs they are defined for (i.e. $\mu = 11$-13 for Δk and $\mu = 6$ for l), the unconnected motif with a bidirectional link ($\mu = 0$) is still not sufficiently well-described without including the industry structure in the model (see Figures C.3 and C.4 in the appendix). Thus, the additional industry structure improves the description of the motif distribution in the model markedly. From Table 6.2 and figures 6.8 and 6.9, we see that, consequentially, the model $\mathcal{G}_{ind,\Delta k,l}$ offers the best description of the motif distribution.

From this observation, we conclude that substitutions of business relations play an important role in the formation of motifs. In particular, our results imply that the representation of the motif distribution in the model improves when business relations are relinked by firms in favor of partners that already exhibit many common business partners.

6.5. Discussion

In this chapter, we have explored different strategies to influence the motif distribution in an evolving network model to better understand the formation of 3-node motifs in the Japanese Business Firm Network. While the original model provides a good description of many topological properties of the real-world network, it overestimates the appearance of densely connected motifs with bidirectional links. To better understand the origin of motif appearances, we have analyzed the topology of individual industry subnetworks in the real-world network and introduced an industry structure to the model. Unlike other industries, the construction sector in Japan follows a unique degree distribution that does not obey a power-law due to the lack of middle-sized firms. The performed role analysis with respect to industries demonstrated that roles with bidirectional links appear predominantly in the manufacturing sector.

We have introduced subnetwork- and industry-specific linking patterns at node creation in the evolving network model. This influences the appearance of sparsely connected motifs within the individual subnetworks. However, the impact on densely connected motifs with bidirectional links is small. To better describe the appearance of densely connected motifs in the real-world network, we modified the merging process of nodes in the model by introducing the Δk-rule. This rule states that nodes that exhibit higher in-degree than out-degree do not merge with nodes that have more outgoing than incoming links. In the context of business networks, this rule is interpreted such that firms which predominantly provide goods or services to other companies do not merge with firms that predominantly exhibit the position of a customer. With the Δk-rule, the frequencies of densely connected motifs with

bidirectional links in the model are reduced, improving the reproduction of the real-world network properties in the model. Furthermore, we have shown that by introducing additional possibilities to reconnect links from a node, preferably to another node with many common neighbors, the deficit of appearances of completely unidirectionally connected motifs in the model is reduced.

Our results provide both empirical and conceptual insights into the mechanisms of motif formation in an evolving network model. We have modified the existing basic model to improve the description of the motif distribution by achieving a better agreement with the real-world business network. In addition, our findings provide information about the selection preferences of business partners in the creation and merging processes of firms. We have established the hypothesis that business firms tend to merge with companies that exhibit similarities in their input/output allocations. This finding provides a starting point for future empirical studies that focus on the merging behavior of firms. In particular, we outline the analysis of the degree distribution of merging partners in the JBFN as an important topic for future research. The second elaborated hypothesis of this study states that companies tend to substitute business relations in favor of companies that share many common business partners. This can be deduced from the observed improvement of the model to better reproduce the appearance of motifs when substitution possibilities are taken into account.

Chapter 7.

Conclusion

In this thesis we have illustrated how methods from complex network theory offer valuable new insights into the trade system of goods and services. By addressing relevant economic questions, we have extended the methodology and introduced new network measures. Thus, we gained new conceptual knowledge regarding complex networks in general, which is in principle applicable to network analyses in a variety of research disciplines.

Network of networks perspective

In chapter 3, we have analyzed the topology of the ITN from a network of networks perspective. By considering community detection algorithms and evaluating different partitions, we concluded that national economies still account for meaningful partitions in the network of industries that trade globally with goods and services. Furthermore, the analysis of the evolution of the network of networks topology suggests that globalization is likely to be an ongoing process that is far from being completed. We have introduced new network measures such as the cross-betweenness fraction and we have extended the concept of the Hamming distance to weighted networks. The cross-betweenness fraction allows for a quantification of the importance of a specific subset of nodes in connecting other subnetworks with each other. With this measure, we have identified subgraphs in the trade network that take an important role as intermediate producers and service providers. Next to some expected results, such as that the world's largest economies also exhibit a large cross-betweenness fraction, we have made some more non-trivial observations. For example, smaller nations like Switzerland take an important role in the global supply chain by connecting the path between two other nations. With respect to the weighted Hamming distance, we have demonstrated that this property serves as a suitable measure to quantify the amount of reorganization of trading patterns. The structure of trade undergoes the largest reorganization in times of crises and shocks. Thus, the impact of the Great Recession that has been triggered by the financial crisis at the end of the first decade of the new millennium can be clearly observed in the time evolution of the weighted Hamming distance.

Interconnectedness between subnetworks

To assess the impact of bilateral trade agreements on the trade structure of the involved economies in chapter 4, we have defined a new measure that is based on the probability that a random walker follows a path from one subnetwork to the other. This method quantifies the interconnectedness between subnetworks and is applicable to flow networks in which the weight of a link can be meaningfully interpreted as a probability that the respective edge is chosen by a random walker. Furthermore, we have developed a method to categorize the short-term evolution of the measure when taking only few time points into account. In this method we utilize both linear and non-linear regression techniques to analyze the trend and consider if the level of the measure has changed significantly. With this methodology, we have analyzed the impact of bilateral trade agreements that have been implemented in the last decades. The findings include that the interconnectedness between two countries that have negotiated an agreement has increased in most cases, as expected from intuitive considerations. However, we have identified notable exceptions from this rule. We have provided empirical evidence that while the USA has acquired new markets with a particular focus on their exports, China's trade agreements play a less important role for its exports and imports. However, from the opposite perspective, i.e. from the view of China's agreement partner's, their exports and imports to China have become more important for them. Thus, China is likely to have gained more political and economic influence on their agreement partners. A similar development from the partners' perspective is not observed in the case of the agreements of the United States.

Functional network of economic performance

We have addressed the question if the local properties in the ITN provide insights into correlations between the sectors' economic performance. In chapter 5 we have discussed methods to identify correlations from short time series and constructed the World Economic Performance Network (WEPN). From a topological characterization of the WEPN, we have observed that industries in the same country are likely to be linked in the WEPN, i.e. positively correlated in their value added. A comparison of the topologies of the ITN and the WEPN has revealed that a large direct trade volume between two industries makes a positive correlation between two industries more likely. However, dependency measures provide an even better indicator towards the existence of a link in the WEPN. While first-order measures of the ITN offer important insights to the structure of the WEPN, we have shown that higher-order measures have to be considered as well, in particular for international trade relations. From these higher-order measures we observed that two industries that trade with a common neighborhood are likely to have similarities in the time evolution of their value added. In the investigation for negative correlations, we found no evidence for a detectable competition effect between two nodes at this level of sectoral aggregation of the ITN. That is, we have found no evidence for the hypothesis that two industries

are likely to be negatively correlated if the topology of the ITN suggests a strong competition for identical consumers or suppliers between the two industries.

Motif formation

Finally, we have investigated the formation of trade relationships at the level of individual business firms. In particular, we have analyzed the formation process of 3-node motifs in a scale-free network topology with an evolving network model. By modifying processes and various mechanisms of the model, we have monitored the impacts of the modifications on the motif distribution. With this approach, we have analyzed the preferences of firms in building their trade relationships with other companies within the Japanese Business Firm Network. We have observed from the simulation model that the choice of business partners in the creation process of newcomer firms only affects the frequency of sparsely connected networks. In order to reproduce the statistical characteristics of the motif distribution of the real-world, the introduction of rules for the merging process as well as possibilities for the substitution of trade relationships are required in the model. Thus, we have developed two hypotheses in the course of this study. The first one states, that if business firms decide for a merging process, they choose a partner with a similar structure of input-output allocations. With the second hypothesis, we propose that business firms tend to substitute business relations in favor of companies that already exhibit many trade relationships with common partners.

Outlook

One apparent task of future work is to look for empirical evidence for the two proposed hypotheses regarding the preferences of Japanese firms in choosing their trade partners that are described above. Although the mechanisms of motif formation have been investigated in the context of the Japanese Business Firm Network, our results can be potentially relevant for other disciplines as well. For example, motifs in biological networks are utilized to describe specific functional roles of the nodes [134]. In particular the relinking process discussed above can offer a starting point for future research on growth models of protein-protein interaction networks [160], in order to better understand the underlying mechanisms that determine the formation process of motifs in these networks.

Our interpretations of the results obtained for the ITN can be further strengthened by a sophisticated mapping of the monetary value of the flows of goods and services to the corresponding physical flows. As different goods perceive price volatilities of different magnitude on short time scales, the absolute values of the monetary flows can draw a distorted picture in terms of the actual amount of traded goods. However, reliable and consistent data of the physical flows on a global scale is difficult to obtain. Thus, such analyses would have to be restricted to specific goods for which data is available.

The presented framework to quantify bilateral trade agreements can be utilized for more detailed case studies. For example, it allows for an investigation of a specific trade agreement to address the question if third parties have been affected positively or negatively by the agreement as well. An intuitive hypothesis would state that a stronger trade interconnectedness between two nations would come at the expense of the trade partners that are not involved in the agreement. Another research opportunity includes an extension of the method and measures to assess the impact of multilateral agreements, such as the North American Free Trade Agreement or the European Union. In these cases, the trade interconnectedness has to be evaluated and compared between all possible country pairs from the set of countries that are involved in the agreement under investigation.

Insights from the comparison between the topologies of the WEPN and the ITN can be utilized to better model the spread of economic shocks in the network of global trade. As the presented analysis is model-independent, the empirical results can be compared with results from the model under investigation. In further research, positive correlations in VAG can then be attributed to either cascading mechanisms through the supply chain or to intrinsic topological properties of the trade network. However, as industries are constantly adapting their operations to their current economic environment, a higher time resolution of the data is desirable for a robust analysis on the spreading of shocks.

Appendix

Appendix A.

Complementary tables on the International Trade Network

Industry classification and countries in Eora

This section provides additional information on the ITN as analyzed in chapters 3-5. Tables A.1 and A.2 list the 26 industrial sectors and 186 countries, respectively, that are included in the utilized Eora database.

Table A.1.: List of sectors that are included in the Eora database

Agriculture
Fishing
Mining and Quarrying
Food & Beverages
Textiles and Wearing Apparel
Wood and Paper
Petroleum, Chemical and Non-Metallic Mineral Products
Metal Products
Electrical and Machinery
Transport Equipment
Other Manufacturing
Recycling
Electricity, Gas and Water
Construction
Maintenance and Repair
Wholesale Trade
Retail Trade
Hotels and Restraurants
Transport
Post and Telecommunications
Financial Intermediation and Business Activities
Public Administration
Education, Health and Other Services
Private Households
Others
Re-export & Re-import

Table A.2.: List of countries that are included in the Eora database. For countries marked with (*), no complete data on the exchange rate and GDP deflator could be obtained. Thus, these countries are excluded from the analysis of the WEPN in chapter 5.

Afghanistan*	Egypt*	Mali	Swaziland
Albania*	El Salvador	Malta	Sweden
Algeria	Eritrea	Mauritania*	Switzerland
Andorra*	Estonia*	Mauritius	Syria*
Angola	Ethiopia	Mexico	TFYR Macedonia*
Antigua	Fiji	Moldova*	Taiwan*
Argentina	Finland	Monaco*	Tajikistan*
Armenia*	France	Mongolia*	Tanzania
Aruba*	French Polynesia*	Montenegro*	Thailand
Australia	Gabon	Morocco	Togo
Austria	Gambia*	Mozambique	Trinidad and Tobago
Azerbaijan*	Gaza Strip*	Myanmar*	Tunisia
Bahamas*	Georgia*	Namibia	Turkey
Bahrain	Germany	Nepal	Turkmenistan*
Bangladesh	Ghana	Netherlands	UAE
Barbados	Greece	Netherlands Antilles*	UK
Belarus*	Greenland*	New Caledonia*	USA
Belgium	Guatemala	New Zealand	Uganda
Belize	Guinea	Nicaragua	Ukraine*
Benin	Guyana	Niger	Uruguay
Bermuda	Haiti*	Nigeria	Uzbekistan*
Bhutan	Honduras	North Korea*	Vanuatu
Bolivia	Hong Kong*	Norway	Venezuela*
Bosnia and Herzegovina*	Hungary	Oman	Viet Nam
Botswana	Iceland	Pakistan	Yemen*
Brazil	India	Panama	Zambia
British Virgin Islands*	Indonesia	Papua New Guinea	Zimbabwe*
Brunei	Iran*	Paraguay	
Bulgaria	Iraq	Peru	
Burkina Faso	Ireland	Philippines	
Burundi	Israel	Poland	
Cambodia*	Italy	Portugal	
Cameroon	Jamaica*	Qatar*	
Canada	Japan	Romania	
Cape Verde	Jordan	Russia	
Cayman Islands*	Kazakhstan*	Rwanda	
Central African Republic	Kenya	Samoa	
Chad	Kuwait	San Marino*	
Chile	Kyrgyzstan*	Sao Tome and Principe*	
China	Laos	Saudi Arabia	
Colombia	Latvia*	Senegal	
Congo*	Lebanon	Serbia*	
Costa Rica	Lesotho	Seychelles	
Cote dIvoire	Liberia	Sierra Leone	
Croatia*	Libya	Singapore	
Cuba*	Liechtenstein*	Slovakia*	
Cyprus	Lithuania*	Slovenia*	
Czech Republic	Luxembourg	Somalia*	
DR Congo*	Macao SAR*	South Africa	
Denmark	Madagascar	South Korea*	
Djibouti	Malawi	Spain	
Dominican Republic	Malaysia	Sri Lanka	
Ecuador	Maldives*	Suriname	

Industry and country specific characterization of roles in the ITN

The following tables complement the results of section 3.4.2 on the assignment of roles of nodes in the ITN. The complete list on industries' strengths in domestic trade (international trade) is depcited in Table A.3 (in Table A.4). The strength of an industry s_q is defined as the sum of the strengths over all nodes that belong to the industry but are of different countries. Tables A.5 and A.6 show the values of the betweenness as defined in equation (3.14) for the 40 countries with the highest scores and for all industrial sectors.

Table A.3.: Internal strength $s_{q;\text{auto}}$ of the industries in the ITN of 2005.

	domestic input $s^{in}_{q;\textbf{auto}}$		**domestic output** $s^{out}_{q;\textbf{auto}}$
%	**industry**	%	**industry**
11.4	Financial Services & Business	23.2	Financial Services & Business
7.4	Electrical & Machinery	8.8	Petroleum, Chemical & Non-Metallic
7.1	Petroleum, Chemical & Non-Metallic	5.4	Transport
6.4	Education, Health & Other Services	5.1	Metal Products
6.2	Construction	5.0	Electrical & Machinery
5.2	Food & Beverages	4.0	Wholesale Trade
4.2	Transport	3.4	Agriculture
3.9	Transport Equipment	3.1	Wood & Paper
3.9	Public Administration	3.1	Electricity, Gas & Water
3.8	Metal Products	2.9	Post & Telecommunications
2.9	Wholesale Trade	2.7	Education, Health & Other Services
2.8	Retail Trade	2.7	Construction
2.4	Hotels & Restraurants	2.5	Food & Beverages
2.4	Agriculture	2.1	Transport Equipment
2.3	Wood & Paper	2.1	Mining & Quarrying
2.1	Electricity, Gas & Water	1.8	Retail Trade
2.0	Textiles & Wearing Apparel	1.2	Textiles & Wearing Apparel
1.9	Post & Telecommunications	0.8	Hotels & Restraurants
1.2	Mining & Quarrying	0.6	Other Manufacturing
1.0	Other Manufacturing	0.4	Maintenance & Repair
0.5	Others	0.3	Public Administration
0.4	Maintenance & Repair	0.3	Others
0.2	Recycling	0.2	Fishing
0.2	Fishing	0.2	Recycling
0.1	Private Households	0.0	Private Households
0.0	Re-export & Re-import	0.0	Re-export & Re-import

Table A.4.: Cross-strength $s_{q;\mathsf{cross}}$ of the industries in the ITN of 2005.

	foreign input $s^{in}_{q;\mathbf{cross}}$		foreign output $s^{out}_{q;\mathbf{cross}}$
%	**industry**	%	**industry**
3.9	Electrical & Machinery	2.9	Electrical & Machinery
3.2	Petroleum, Chemical & Non-Metallic	2.7	Petroleum, Chemical & Non-Metallic
1.6	Metal Products	1.4	Transport Equipment
1.5	Financial Services & Business	1.3	Re-export & Re-import
1.3	Mining & Quarrying	1.2	Financial Services & Business
1.0	Re-export & Re-import	1.0	Metal Products
1.0	Transport Equipment	0.9	Construction
0.9	Transport	0.8	Food & Beverages
0.8	Wood & Paper	0.8	Transport
0.6	Food & Beverages	0.7	Education, Health & Other Services
0.6	Textiles & Wearing Apparel	0.6	Wood & Paper
0.4	Agriculture	0.6	Textiles & Wearing Apparel
0.4	Wholesale Trade	0.5	Public Administration
0.2	Post & Telecommunications	0.4	Wholesale Trade
0.2	Education, Health & Other Services	0.3	Electricity, Gas & Water
0.2	Other Manufacturing	0.3	Post & Telecommunications
0.1	Retail Trade	0.3	Retail Trade
0.1	Hotels & Restraurants	0.3	Other Manufacturing
0.0	Recycling	0.3	Hotels & Restraurants
0.0	Construction	0.2	Agriculture
0.0	Fishing	0.2	Mining & Quarrying
0.0	Electricity, Gas & Water	0.1	Maintenance & Repair
0.0	Maintenance & Repair	0.0	Others
0.0	Public Administration	0.0	Recycling
0.0	Others	0.0	Fishing
0.0	Private Households	0.0	Private Households

Table A.5.: Countries with the largest average cross-betweenness as defined in equation (3.14) in the ITN of 2005, aggregated by country.

National partition $\mathcal{C}_c$		**Sectoral partition $\mathcal{C}_s$**	
$\sum \langle b_i \rangle$	**country**	$\sum \langle b_i \rangle$	**country**
0.145	Germany	0.114	USA
0.103	USA	0.074	Germany
0.089	Switzerland	0.064	China
0.079	UK	0.038	France
0.066	China	0.036	Netherlands
0.061	Netherlands	0.032	Italy
0.050	Japan	0.030	Belgium
0.049	Italy	0.029	UK
0.043	France	0.025	Japan
0.041	Belgium	0.020	South Africa
0.022	Sweden	0.019	Russia
0.017	Denmark	0.015	India
0.017	Canada	0.012	Switzerland
0.016	Australia	0.011	Spain
0.015	Hong Kong	0.010	South Korea
0.014	Malaysia	0.009	Mauritania
0.014	South Korea	0.008	Burundi
0.013	India	0.008	Australia
0.013	Spain	0.008	Singapore
0.012	Russia	0.008	Eritrea
0.012	Taiwan	0.007	Gaza Strip
0.011	Indonesia	0.007	Rwanda
0.009	Singapore	0.007	Denmark
0.009	Ireland	0.006	Afghanistan
0.009	Finland	0.006	Vanuatu
0.008	Austria	0.006	Brazil
0.007	Thailand	0.006	Central African Republic
0.006	Norway	0.006	Lesotho
0.006	Philippines	0.006	Chad
0.005	Brazil	0.006	Cape Verde
0.004	South Africa	0.005	Jordan
0.003	Turkey	0.005	Seychelles
0.003	Hungary	0.005	Bhutan
0.003	Poland	0.005	Burkina Faso
0.003	Mexico	0.005	Samoa
0.003	Czech Republic	0.005	Malaysia
0.002	New Zealand	0.004	Laos
0.002	Ukraine	0.004	Canada
0.002	Saudi Arabia	0.004	Mali
0.002	Viet Nam	0.004	Thailand
.	.	.	.

Table A.6.: Average cross-betweenness as defined in equation (3.14) in the ITN of 2005, aggregated by industry.

National partition $\mathcal{C}_c$		**Sectoral partition $\mathcal{C}_s$**	
$\sum \langle b_i \rangle$	**industries**	$\sum \langle b_i \rangle$	**industries**
0.178	Re-export & Re-import	0.193	Finacial Services & Business Activities
0.153	Petroleum	0.146	Petroleum
0.152	Finacial Services & Business Activities	0.115	Electrical and Machinery
0.119	Electrical and Machinery	0.107	Re-export & Re-import
0.060	Metal Products	0.098	Transport
0.047	Transport	0.060	Food & Beverages
0.043	Wood and Paper	0.045	Metal Products
0.038	Education	0.031	Education
0.035	Food & Beverages	0.027	Mining and Quarrying
0.028	Textiles and Wearing Apparel	0.025	Agriculture
0.026	Wholesale Trade	0.024	Wood and Paper
0.023	Transport Equipment	0.022	Hotels and Restraurants
0.020	Retail Trade	0.018	Wholesale Trade
0.019	Hotels and Restraurants	0.017	Transport Equipment
0.019	Other Manufacturing	0.016	Textiles and Wearing Apparel
0.011	Post and Telecommunications	0.013	Post and Telecommunications
0.011	Agriculture	0.010	Construction
0.006	Mining and Quarrying	0.010	Other Manufacturing
0.004	Construction	0.009	Public Administration
0.003	Electricity	0.008	Retail Trade
0.003	Public Administration	0.003	Electricity
0.001	Recycling	0.001	Recycling
0.000	Maintenance and Repair	0.001	Fishing
0.000	Fishing	0.000	Others
0.000	Others	0.000	Maintenance and Repair
0.000	Private Households	0.000	Private Households

Appendix B.

Supplementary information: The effects of bilateral trade agreements

Sensitivity analysis on the influence of the maximal path length

Here, we investigate the influence of the parameter λ_{max} on the results presented in chapter 4. The countries' average BTA impacted indices for the maximal path lengths $\lambda_{max} = 1$ and $\lambda_{max} = 10$ are shown for their inputs in Figure B.1 and for their outputs in Figure B.2, respectively. We observe in many countries a tendency towards smaller values of the input BTA impact indices Π_x^{in} with increasing maximal path length. This tendency applies to, for example, Europe, Australia, Algeria, Kazakhstan and Central America. The negative tendency is less pronounced for the output BTA impact indices Π_x^{out}. This can be explained by the arguments discussed in section 4.2.1: The contribution of national loops in Π_x^{in} increase with larger λ_{max}, while the value of Π_x^{out} is already close to the saturation level at $\lambda_{max} = 6$ (see Figure 4.3). Thus, the trend in the ITN towards more international trade, which has a negative influence on $\Pi^{\bullet}$, is more pronounced in Π^{in} than in Π^{out}.

For a more detailed view on the trade profile of China, we illustrate the results for the choices of $\lambda_{max} = 1$ in Figure B.4 and for $\lambda_{max} = 10$ in Figure B.5. We observe that the trade agreements of China with New Zealand and Hong Kong follow the general tendency towards a lower input BTA impact index with increasing maximal path length (see top right panels in the figures). However, in the input TI of China to Pakistan, a higher maximal path length increases the BTA impact index. The corresponding time series of the TI are shown in Figure B.3. It can be seen that the higher BTA impact index is attributed to a changing behavior of the TI^{in} in 2009 with increasing λ_{max}. In this year, the Great Recession triggered by the financial crisis caused a decline in international trade, interrupting the general trend in this year (see section 3.4.3). Thus, in this exceptional year, higher probabilities for national loops are likely to be observed as compared to the previous and following years. This exception is responsible for the increase of the input BTA impact index of China to Pakistan for increasing λ_{max}.

In the other panels of Figures B.4 and B.5, we see that the TI of the partners to China is stable for both the output TI (bottom left panel) and the input TI (bottom right panel). For the output TI of China to its partners (top left panel), no negative

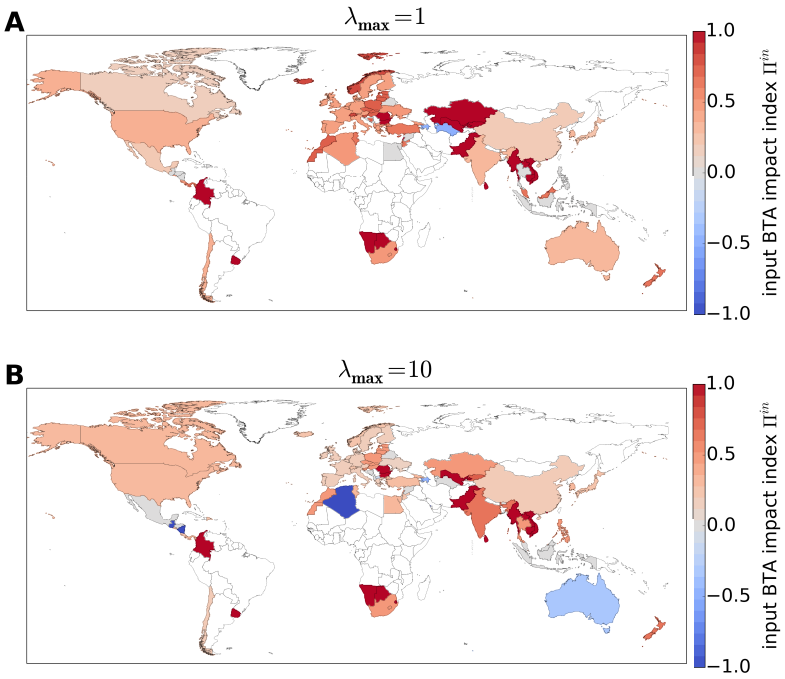


Figure B.1.: The countries' average input BTA impact indices Π^{in} for (A) $\lambda_{max} = 1$ and (B) $\lambda_{max} = 10$.

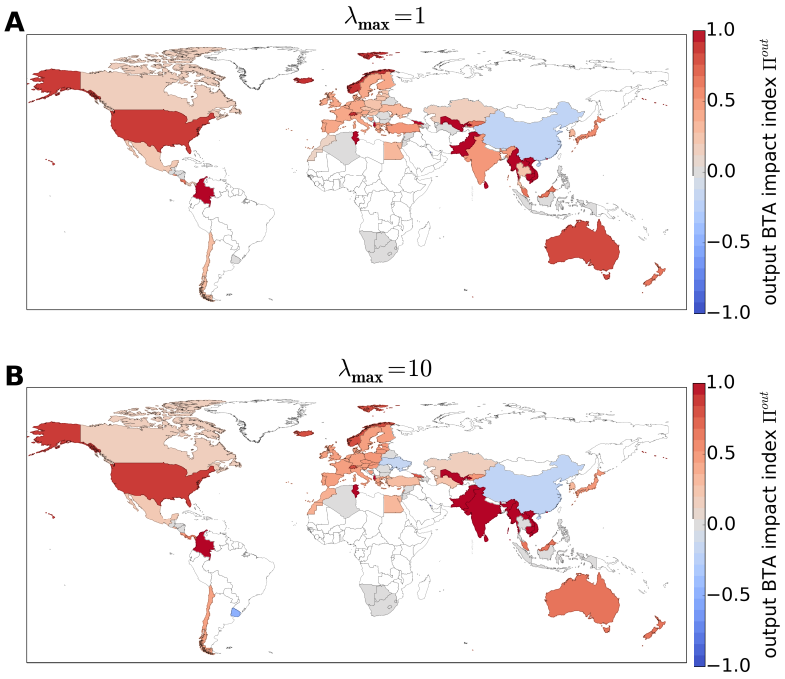


Figure B.2.: The countries' average output BTA impact indices Π^{in} for (A) $\lambda_{max} = 1$ and (B) $\lambda_{max} = 10$.

tendency in the impact index with higher λ_{max} can be observed. This coincides with the arguments discussed above. The trade profiles of the USA for different maximal path lengths are shown in Figures B.6 ($\lambda_{max} = 1$) and B.7 ($\lambda_{max} = 10$).

From these findings, we conclude that higher-order paths have an impact on the TI and should be considered in the analysis on the impact of trade agreements. The main conclusions that we draw from the results in chapter 4, however, are robust with the choice of reasonable values of λ_{max}.

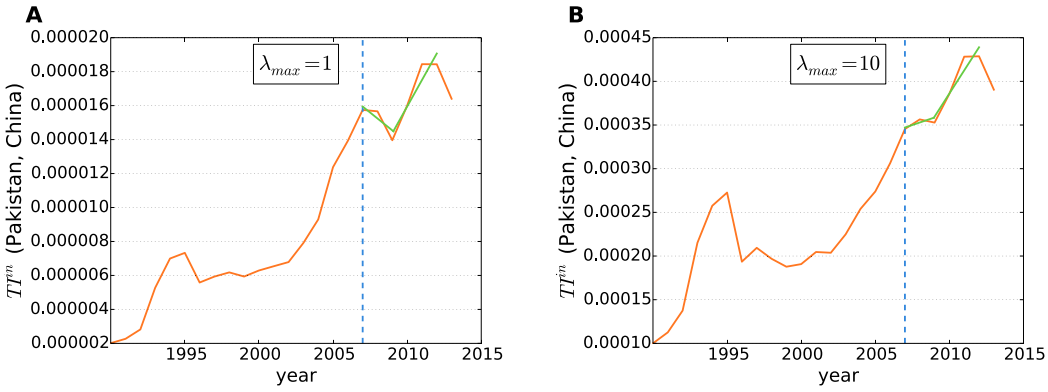


Figure B.3.: The input TI of China to Pakistan for (A) $\lambda_{max} = 1$ and (B) $\lambda_{max} = 10$. The year in which the BTA comes into effect is indicated by the blue vertical line. The regression model selected by the AIC criterion and the corresponding maximum likelihood fit is displayed by the green line.

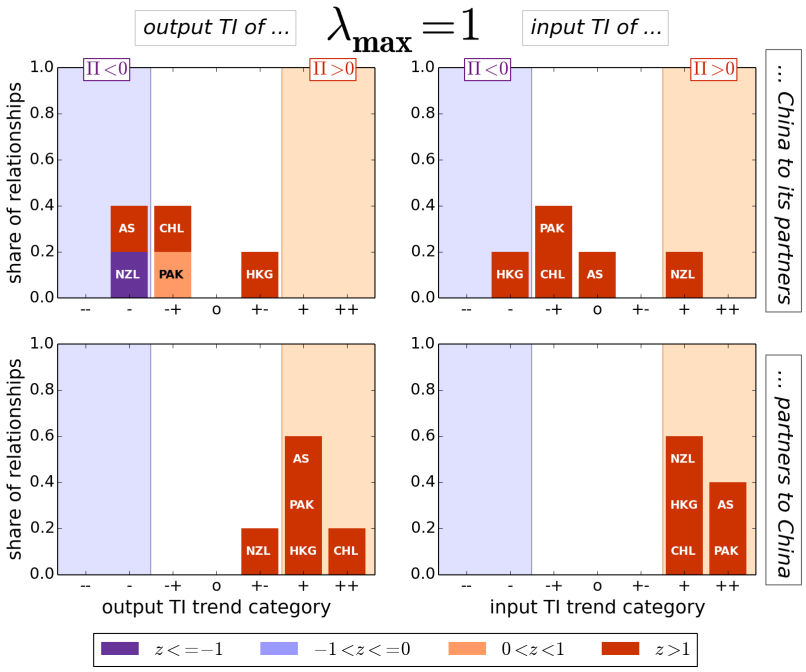


Figure B.4.: The trade profile of China for $\lambda_{max} = 1$, with definitions as in Figure 4.8.

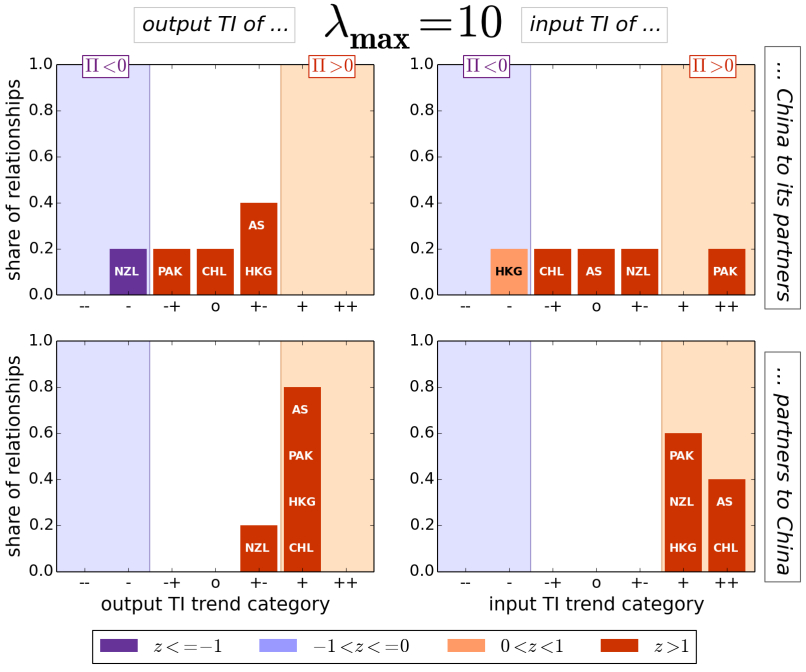


Figure B.5.: The trade profile of China for $\lambda_{max} = 10$, with definitions as in Figure 4.8.

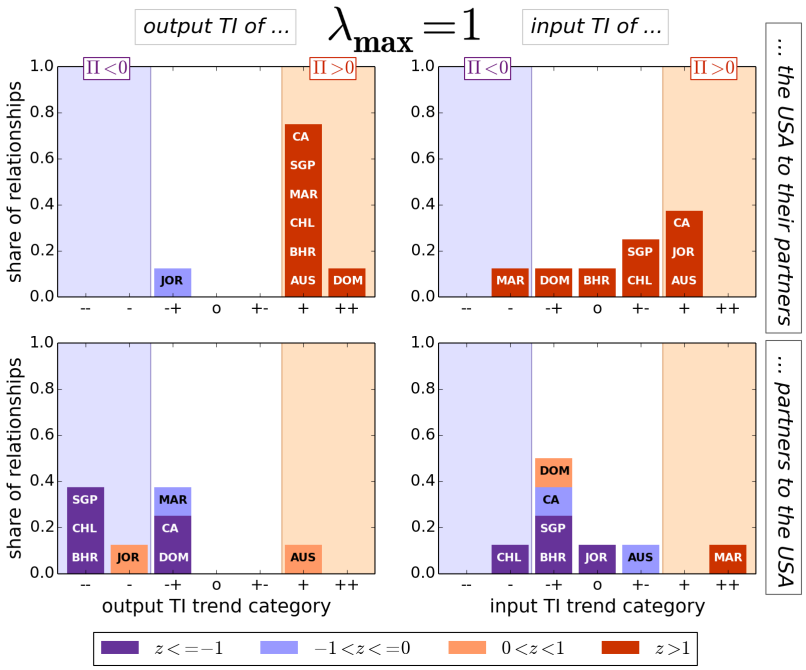


Figure B.6.: The trade profile of the USA for $\lambda_{max} = 1$, with definitions as in Figure 4.8.

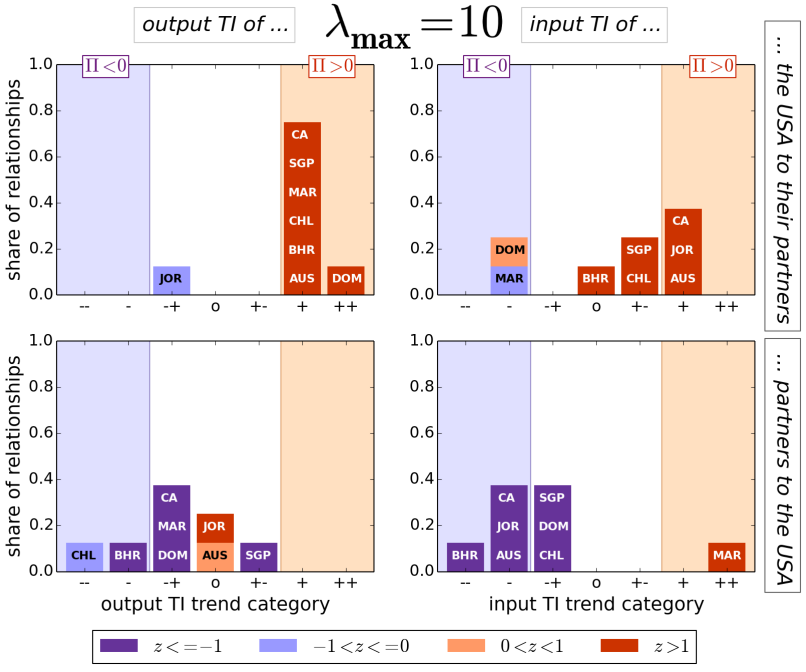


Figure B.7.: The trade profile of the USA for $\lambda_{max} = 10$, with definitions as in Figure 4.8.

List of all bilateral trade agreements between 1995 and 2008

The 107 bilateral trade agreements that were notified to the WTO and that have been negotiated with a date of entry into force between 1995 and 2008 are listed in table B.1. Bilateral trade agreements in which one party is a regional trade block are included in the analysis. In particular the following regional trade blocks have negotiated a bilateral agreement with another party:

ASEAN	Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, Philippines, Singapore, Thailand, Viet Nam
CACM	Costa Rica, El Salvador, Guatemala, Honduras, Nicaragua
EC15	Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, UK
EC25	all of the EC15 + Cyprus, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, Slovenia
EC27	all of the EC25 + Bulgaria, Romania
EFTA	Iceland, Liechtenstein, Norway, Switzerland
SACU	Botswana, Lesotho, Namibia, South Africa, Swaziland

Agreement partners		**Date**	**Av. path length →**	**Av. path length ←**
Albania	Turkey	2008	3.60	2.94
Armenia	Georgia	1998	0.00	0.00
Armenia	Kazakhstan	2001	0.00	0.00
Armenia	Kyrgyzstan	1995	0.00	0.00
Armenia	Moldova	1995	0.00	0.00
Armenia	Turkmenistan	1996	3.61	0.00
Armenia	Ukraine	1996	3.67	0.00
Australia	Singapore	2003	1.97	1.83
Australia	Thailand	2005	2.15	1.90
Australia	USA	2005	1.82	1.62
Azerbaijan	Georgia	1996	3.09	0.00
Azerbaijan	Ukraine	1996	3.76	2.93
Bahrain	USA	2006	2.69	2.36
Belarus	Ukraine	2006	3.00	2.11
Bhutan	India	2006	3.65	0.00
Bosnia and Herzegovina	Turkey	2003	2.52	3.38
Brunei	Japan	2008	2.59	2.12
Canada	Chile	1997	2.08	2.16
Canada	Costa Rica	2002	2.79	2.64
Canada	Israel	1997	2.12	2.07
Chile	China	2006	2.08	2.03
Chile	Costa Rica	2002	2.93	0.00
Chile	El Salvador	2002	0.00	0.00
Chile	Honduras	2008	0.00	0.00
Chile	Japan	2007	1.97	2.16
Chile	Mexico	1999	2.01	2.12
Chile	Panama	2008	3.05	0.00
Chile	South Korea	2004	2.13	2.28
Chile	USA	2004	1.99	1.86
China	Hong Kong	2003	1.41	1.65
China	New Zealand	2008	1.94	2.02

China	Pakistan	2007	2.13	2.33
Colombia	Mexico	1995	2.20	2.01
Costa Rica	Panama	2008	2.53	2.81
Dominican Republic	USA	2006	2.04	1.92
Egypt	Turkey	2007	2.52	2.35
El Salvador	Panama	2003	3.33	2.80
Georgia	Kazakhstan	1999	2.25	2.23
Georgia	Turkey	2008	2.79	2.89
Georgia	Turkmenistan	2000	0.00	2.38
Georgia	Ukraine	1996	3.21	3.61
India	Singapore	2005	1.98	1.91
India	Sri Lanka	2001	2.09	2.82
Indonesia	Japan	2008	1.74	1.83
Israel	Mexico	2000	2.28	2.96
Israel	Turkey	1997	2.19	2.17
Japan	Malaysia	2006	1.69	1.76
Japan	Mexico	2005	1.87	1.96
Japan	Philippines	2008	1.79	1.83
Japan	Singapore	2002	1.78	1.84
Japan	Thailand	2007	1.88	1.83
Jordan	Singapore	2005	3.05	0.00
Jordan	USA	2001	3.29	2.27
Kazakhstan	Kyrgyzstan	1995	2.89	2.69
Kazakhstan	Ukraine	1998	2.77	2.48
Kyrgyzstan	Moldova	1996	0.00	0.00
Kyrgyzstan	Ukraine	1998	0.00	0.00
Kyrgyzstan	Uzbekistan	1998	3.88	3.88
Malaysia	Pakistan	2008	2.42	2.84
Mexico	Uruguay	2004	3.08	2.72
Moldova	Ukraine	2005	0.00	0.00
Morocco	Turkey	2006	2.78	2.92
Morocco	USA	2006	2.38	2.47
New Zealand	Singapore	2001	2.31	2.32
New Zealand	Thailand	2005	2.50	2.31
Pakistan	Sri Lanka	2005	0.00	0.00
Panama	Singapore	2006	3.46	0.00
Russia	Serbia	2006	0.00	0.00
Singapore	South Korea	2006	1.99	2.04
Singapore	USA	2004	1.82	1.78
Syria	Turkey	2007	2.45	2.68
Macedonia	Turkey	2000	3.04	2.81
Macedonia	Ukraine	2001	0.00	3.79
Tajikistan	Ukraine	2002	0.00	0.00
Tunisia	Turkey	2005	3.00	2.73
Turkmenistan	Ukraine	1995	2.96	2.40
Ukraine	Uzbekistan	1996	0.00	0.00
Albania	EC25	2006	2.56	2.18
Algeria	EC25	2005	1.93	1.86
Bosnia and Herzegovina	ec27	2008	2.48	2.22
Chile	EC15	2003	1.82	1.96
Chile	EFTA	2004	2.52	2.29
China	ASEAN	2005	1.61	1.73
Dominican Republic	CACM	2001	0.00	3.02
Egypt	EC25	2004	1.99	1.94

Egypt	EFTA	2007	2.90	2.33
Israel	EC15	2000	1.71	1.87
Jordan	EC15	2002	2.72	2.08
Jordan	EFTA	2002	0.00	2.82
Lebanon	EC15	2003	2.54	2.03
Lebanon	EFTA	2007	2.98	2.52
Mexico	EC15	2000	1.77	1.76
Mexico	EFTA	2001	2.58	2.03
Montenegro	ec27	2008	0.00	0.00
Morocco	EC15	2000	1.91	1.93
Morocco	EFTA	1999	2.82	2.68
San Marino	EC15	2002	0.00	0.00
Singapore	EFTA	2003	2.48	2.06
South Africa	EC15	2000	1.62	1.77
South Korea	EFTA	2006	2.10	2.00
Macedonia	EC15	2001	2.13	2.14
Macedonia	EFTA	2002	3.47	2.84
Tunisia	EC15	1998	1.85	1.89
Tunisia	EFTA	2005	3.28	2.54
Turkey	EC15	1996	1.63	1.72
USA	CACM	2006	1.86	1.96
EFTA	SACU	2008	2.08	2.15

Table B.1.: Table summarizing all BTAs recorded by the World Trade Organization that came into force between 1995 and 2008. Parties can be individual countries or trade blocs. Trade blocs that have negotiated a BTA in this time period are ASEAN (Association of Southeast Asian Nations), CACM (Central American Common Market), EC (European Community), EFTA (European Free Trade Association) and SACU (Southern African Customs Union). The average path lengths between the corresponding subnetworks are listed with the arrow depicting the direction of the flow of goods. For example the average path length for the flow of goods from Albania to Turkey is 3.6, while taking a value of 2.94 from Turkey to Albania.

Appendix C.

Motif distributions for variants to model the Japanese Business Firm Network

The variants described in chapter 6 introduce new parameters to the evolving network model that reproduces statistical characteristics of the JBFN. In particular, with the introduction of relinking possibilities, reasonable values for the parameter δ have to be determined. The paramater describes the share of relinking processes in the evolution of the model. Figures C.1 and C.2 show the influences on the motif distribution for model variants $\mathcal{G}_{ind,\Delta k,l}$ with the industry structure and the Δk-rule for the merging process included. The corresponding penalty values for different values of δ are listed in table C.1. We observe that for $\delta = 0.3$ the lowest penalty score is obtained. From the figures, we see that the relinking possibilities in the model mostly affect the frequency of motif patterns $\mu = 6$-8. A high value of δ improves the description of pattern $\mu = 6$ but increases the penalty value for patterns $\mu = 7, 8$. Thus, $\delta = 0.3$ is identified as a reasonable value with a good ballance of these two opposing effects.

Table C.1.: Sum of the penalty values for all motif patterns in the model variant $\mathcal{G}_{ind,\Delta k,l}$ for different values of the relinking probability δ.

model	$\mathcal{G}_{\delta=0.1}$	$\mathcal{G}_{\delta=0.2}$	$\mathcal{G}_{\delta=0.3}$	$\mathcal{G}_{\delta=0.4}$
$\lvert\Delta P(\mathcal{G}_r, \mathcal{G}'_\bullet)\rvert$	15^{+7}_{-4}	15^{+7}_{-3}	14^{+5}_{-3}	15^{+6}_{-4}

Figures C.3 and C.4 show the motif distributions for different combinations of the various model modifications. We observe, that in all model variants that do not include an industry structure ($\mathcal{G}_{\Delta k}$, $\mathcal{G}_l$ and $\mathcal{G}_{\Delta k,l}$) the appearances of sparsely connected motifs are not well described by the model. On the other hand, models without the Δk-rule ($\mathcal{G}_l$ and $\mathcal{G}_{ind,l}$) exhibit an overshoot of bidirectional links (motifs $\mu = 11$-13).

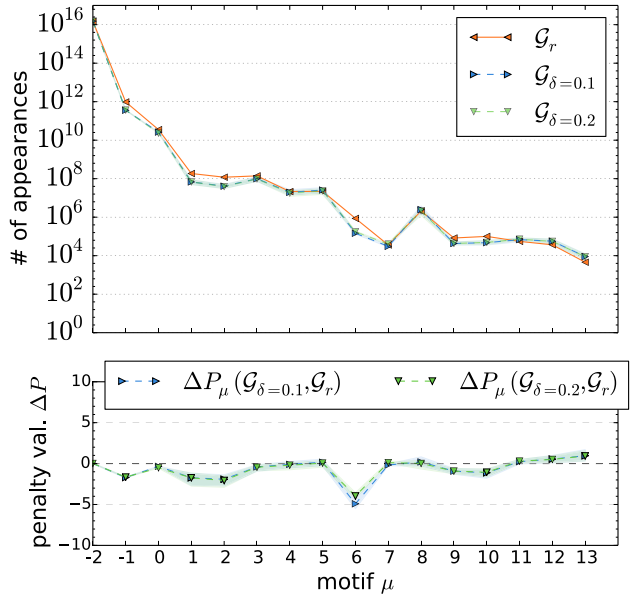


Figure C.1.: Motif distributions as in Figure 6.8 for the model variant $\mathcal{G}_{ind,\Delta k,l}$ for the relinking probabilities $\delta = 0.1$ (light blue) and $\delta = 0.2$ (green).

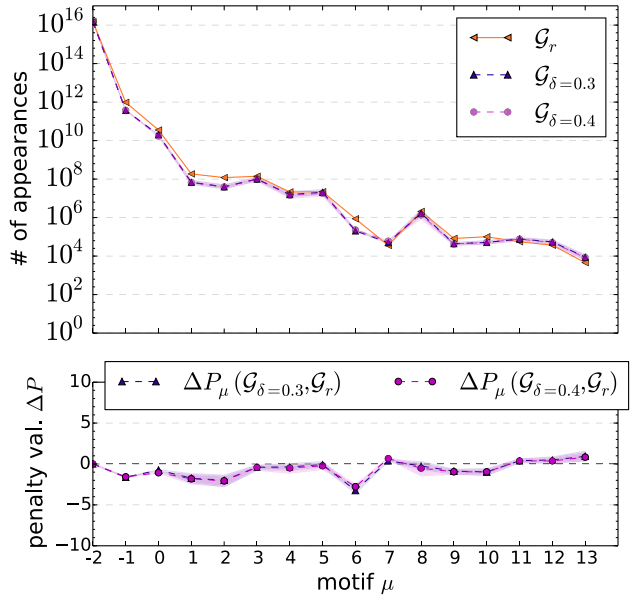


Figure C.2.: Motif distributions as in Figure 6.8 for the model variant $\mathcal{G}_{ind,\Delta k,l}$ for the relinking probabilities $\delta = 0.3$ (dark blue) and $\delta = 0.4$ (magenta).

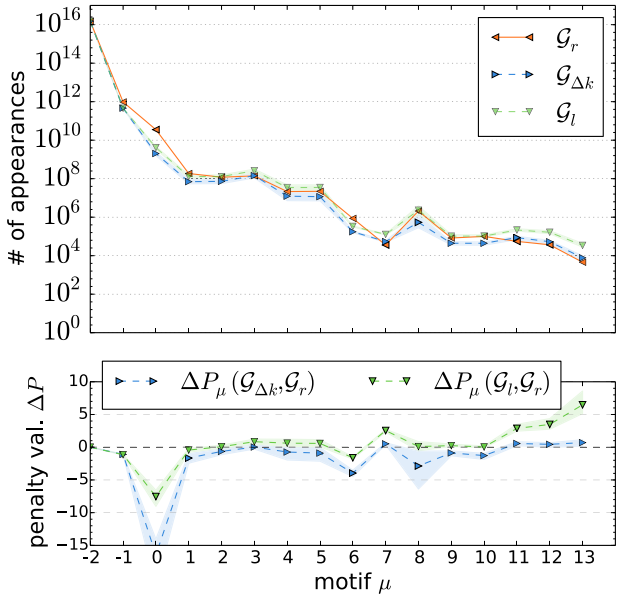


Figure C.3.: Motif distributions as in Figure 6.8 for the model variant $\mathcal{G}_{\Delta k}$ (including only the Δk-rule, light blue) and the variant $\mathcal{G}_l$ (including only relinking possibilities, green).

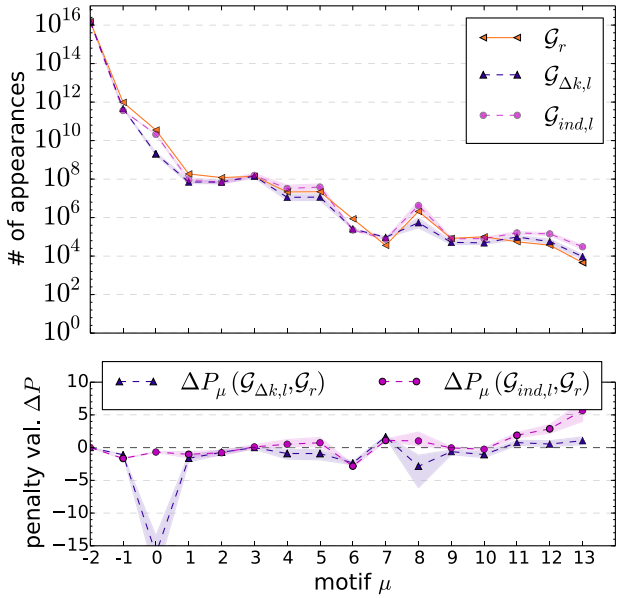


Figure C.4.: Motif distributions as in Figure 6.8 for the model variant $\mathcal{G}_{\Delta k,l}$ (dark blue) and the variant $\mathcal{G}_{ind,l}$ (magenta).

Bibliography

[1] D. A. Irwin. *Against the tide: an intellectual history of free trade*. Princeton, NJ: Princeton University Press, 1996.

[2] P. Wolfowitz. *Trade: The Missing Link to Opportunity*. Remarks at National Press Club, Washington D.C. `https://openknowledge.worldbank.org/handle/10986/24886` (Accessed 09 October 2017). 2005.

[3] J. B. Jensen, D. P. Quinn, and S. Weymouth. "Winners and losers in international trade: The effects on US presidential voting". In: *International Organization* 71(3) (2017), pp. 423–457.

[4] N. G. Mankiw. *Principles of microeconomics*. 8th edition. Stamford, 2017.

[5] O. Blanchard. *Macroeconomics*. Upper Saddle River, NJ: Prentice Hall, 1997.

[6] P. Kennedy. *A guide to econometrics*. 6th edition. Malden, Massachusetts: Blackwell, 2010.

[7] S. Boccalettia, V. Latorab, Y. Morenod, M. Chavezf, and D.-U. Hwanga. "Complex networks: structure & dynamics". In: *Physics Reports* 424 (2006), pp. 175–308.

[8] M. E. J. Newman. *Networks: An introduction*. 1st edition. Oxford University Press, USA, 2010.

[9] L. Euler. "Solutio problematis ad geometriam situs pertinentis". In: *Commentarii academiae scientiarum Petropolitanae* 8 (1741), pp. 128–140.

[10] S. Wasserman and K. Faust. *Social network analysis: methods and applications*. Cambridge: Cambridge University Press, 1994.

[11] R. N. Mantegna and H. E. Stanley. *An introduction to econophysics: correlations and complexity in finance*. Cambridge: Cambridge University Press, 2007.

[12] A. Wagner. "The yeast protein interaction network evolves rapidly and contains few redundant duplicate genes". In: *Molecular Biology and Evolution* 18(7) (2001), pp. 1283–1292.

[13] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, and E. D. Gilles. "Metabolic network structure determines key aspects of functionality and regulation". In: *Nature* 420(6912) (2002), pp. 190–193.

[14] M. Faloutsos, P. Faloutsos, and C. Faloutsos. "On power law relationships of the internet topology". In: *ACM/SIGCOMM Computer Communication Review* 29(4) (1999), pp. 251–262.

[15] F. Varela, J.-P. Lachaux, E. Rodriguez, and J. Martinerie. "The brainweb: Phase synchronization and large-scale integration". In: *Nature Reviews Neuroscience* 2(4) (2001), pp. 229–239.

[16] J. F. Donges, I. Petrova, A. Loew, N. Marwan, and J. Kurths. "How complex climate networks complement eigen techniques for the statistical analysis of climatological data". In: *Climate Dynamics* (2015), pp. 1–18.

[17] C. A. Hidalgo, B. Klinger, A.-L. Barabási, and R. Hausmann. "The product space conditions the development of nations". In: *Science* 317(5937) (2007), pp. 482–487.

[18] E. Eisenberg and E. Y. Levanon. "Preferential Attachment in the Protein Network Evolution". In: *Physical Review Letters* 91 (2003), p. 138701.

[19] H. Hethcote. "The Mathematics of Infectious Diseases". In: *SIAM Review* 42.4 (2000), pp. 599–653.

[20] M. E. J. Newman, S. Forrest, and J. Balthrop. "Email networks and the spread of computer viruses". In: *Physical Review E* 66 (2002), p. 035101.

[21] A. Serrano and M. Boguna. "Topology of the World Trade Web". In: *Physical Review E* 68 (2003), 015101(R).

[22] D. Garlaschelli and M. I. Loffredo. "Structure and evolution of the world trade network". In: *Physica A* 255 (2005), pp. 138–144.

[23] G. Fagiolo, S. Schiavo, and J. Reyes. "On the topological properties of the World Trade Web: a weighted network analysis". In: *Physica A* 387(15) (2008), pp. 3868–3873.

[24] R. Kali and J. Reyes. "The architecture of globalization: A network approach to international economic integration". In: *Journal of International Business Studies* 38(4) (2007), pp. 595–620.

[25] K. Bagwell and R. W. Staiger. "Multilateral trade negotiations, bilateral opportunism and the rules of GATT/WTO". In: *Journal of International Economics* 63 (2004), pp. 1–29.

[26] R. Bierkandt, L. Wenz, S. Willner, and A. Levermann. "Acclimate: a model for economic damage propagation. Part 1: basic formulation of damage transfer within a global supply network and damage conserving dynamics". In: *Environment Systems and Decisions* 43(4) (2014), pp. 507–524.

[27] L. Wenz, S. Willner, R. Bierkandt, and A. Levermann. "Acclimate - a model for economic damage propagation. Part II: a dynamic formulation of the backward effects of disaster-induced production failures in the global supply network". In: *Environment Systems and Decisions* 34(4) (2014), pp. 525–539.

[28] A. Garas, P. Argyrakis, C. Rozenblat, M. Tomassini, and S. Havlin. "Worldwide spreading of economic crisis". In: *New Journal of Physics* 12 (2010), p. 113043.

[29] R. W. Hamming. "Error detecting and error correcting codes". In: *Bell System Technical Journal* 29 (1950), pp. 147–160.

[30] R. Bems, R. C. Johnson, and K.-M. Yi. "Demand spillovers and the collapse of trade in the global recession". In: *IMF Working Paper* WP/10/142 (2010).

[31] X. Li, Y. Y. Jin, and G. Chen. "Complexity and synchronization of the World Trade Web". In: *Physica A* 328 (2003), pp. 287–296.

[32] M. Lenzen, K. Kanemoto, D. Moran, and A. Geschke. "Mapping the structure of the world economy". In: *Environmental Science & Technology* 46(15) (2012), pp. 8364–8381.

[33] M. Lenzen, D. Moran, K. Kanemoto, and A. Geschke. "Building Eora: A global multi-regional input-output database at high country and sector resolution". In: *Economic Systems Research* 25(1) (2013), pp. 20–49.

[34] W. Miura, H. Takayasu, and M. Takayasu. "Effect of coagulation of nodes in an evolving complex network". In: *Physical Review Letters* 108(16) (2012), p. 168701.

[35] J. Murray and M. Lenzen. *The Sustainability practitioner's guide to multi-regional input-output analysis*. Champaign: Common Ground Publishing LLC, 2013.

[36] W. Leontief. *The structure of American economy: 1919 - 1939*. New York: Oxford University Press, 1953.

[37] European Communities. *Eurostat manual of Supply, Use and Input-Output Tables*. Luxembourg: Office for Official Publications of the European Communities, 2008.

[38] A. Rose and W. Miernyk. "Input-output analysis: the first fifty years". In: *Economic Systems Research* 1(2) (1989), pp. 229–272.

[39] European Communities, International Monetary Fund, OECD, United Nations and World Bank. *System of National Accounts 2008*. New York, 2009.

[40] T. Wiedmann. "A review of recent multi-region input-output models used for consumption-based emission and resource accounting". In: *Ecological Economics* 69 (2009), pp. 211–222.

[41] R. V. Donner, E. Hernández-García, and E. Ser-Giacomi. "Introduction to Focus Issue: Complex network perspectives on flow systems". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 27(3) (2017), p. 035601.

[42] V. Rossi, E. Ser-Giacomi, C. López, and E. Hernández-García. "Hydrodynamic provinces and oceanic connectivity from a transport network help designing marine reserves". In: *Geophysical Research Letters* 41(8) (2014), pp. 2883–2891.

[43] L. Wenz and A. Levermann. "Enhanced economic connectivity to foster heat stress-related losses". In: *Science Advances* 2(6) (2016), e1501026.

[44] J. Saramäki, M. Kivelä, J. Onnela, K. Kaski, and J. Kertész. "Generalizations of the clustering coefficient to weighted complex networks". In: *Physical Review E* 75(2) (2007), p. 027105.

[45] G. Fagiolo. "Clustering in complex directed networks". In: *Physical Review E* 76 (2007), p. 026107.

[46] J. Onnela, J. Saramäki, J. Kertész, and K. Kaski. "Intensity and coherence of motifs in weighted complex networks". In: *Physical Review E* 71(6) (2005), p. 065103.

[47] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. "Network motifs: Simple building blocks of complex networks". In: *Science* 298(5594) (2002), pp. 824–827.

[48] N. Kashtan, S. Itzkovitz, R. Milo, and U. Alon. "Topological generalizations of network motifs". In: *Physical Review E* 70(3) (2004), p. 031909.

[49] P. Schultz, J. Heitzig, and J. Kurths. "Detours around basin stability in power networks". In: *New Journal of Physics* 16(12) (2014), p. 125001.

[50] L. C. Freeman. "A set of measures of centrality based on betweenness". In: *Sociometry* 40(1) (1977), pp. 35–41.

[51] A. Radebach, R. V. Donner, J. Runge, J. F. Donges, and J. Kurths. "Disentangling different types of El Nino episodes by evolving climate network analysis". In: *Physical Review E* 88 (2013), p. 052807.

[52] K. Yamasaki, A. Gozolchiani, and S. Havlin. "Climate networks around the globe are significantly affected by El Nino". In: *Physical Review Letters* 100 (2008), p. 228501.

[53] A. Gozolchiani, K. Yamasaki, O. Gazit, and S. Havlin. "Pattern of climate network blinking links follows El Nino events". In: *Europhysics Letters* 83 (2008), p. 28005.

[54] P. Erdős and A. Rényi. "On random graphs". In: *Publicationes Mathematicae (Debrecen)* 6 (1959), pp. 290–297.

[55] S. Fortunato. "Community detection in graphs". In: *Physics Reports* 486 (2010), pp. 75–174.

[56] J. F. Donges, H. C. H. Schultz, N. Marwan, Y. Zou, and J. Kurths. "Investigating the topology of interacting networks". In: *European Physical Journal B* 84 (2011), pp. 635–651.

[57] G. D'Agostino and A. Scala. *Networks of networks: The last frontier of complexity*. Springer International Publishing Switzerland, 2014.

[58] S. Vitali, J. B. Glattfelder, and S. Battiston. "The network of global corporate control". In: *PLoS ONE* 6(10) (2011), e25995.

[59] M. E. J. Newman and M. Girvan. "Finding and evaluating community structure in networks". In: *Physical Review E* 69(2) (2004), p. 026113.

[60] F. D. Malliaros and M. Vazirgiannis. "Clustering and community detection in directed networks: A survey". In: *Physics Reports* 533(4) (2013), pp. 95–142.

[61] A. Arenas, J. Duch, A. Fernández, and S. Gómez. "Size reduction of complex networks preserving modularity". In: *New Journal of Physics* 9(6) (2007), p. 176.

[62] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. "Random graphs with arbitrary degree distributions and their applications". In: *Physical Review E* 64 (2001), p. 026118.

[63] Y. Kim, S. W. Son, and H. Jeong. "Finding communities in directed networks". In: *Physical Review E* 81 (2010), p. 016103.

[64] M. Meila. "Comparing clusterings - an information based distance". In: *Journal of Multivariate Analysis* 98 (2007), pp. 873–895.

[65] WTO Secretariat. *World Trade Report 2014*. Geneva: World Trade Organization, 2014.

[66] M. Karpiarz, P. Fronczak, and A. Fronczak. "International trade network: Fractal properties and globalization puzzle". In: *Physical Review Letters* 113 (2014), p. 248701.

[67] K. Bhattacharya, G. Mukherjee, J. Sarämaki, K. Kaski, and S. S. Manna. "The international trade network: weighted network analysis and modeling". In: *Journal of Statistical Mechanics* 2008(2) (2008), P02002.

[68] G. Fagiolo, J. Reyes, and S. Schiavo. "The evolution of the world trade web: a weighted-network analysis". In: *Journal of Evolutionary Economics* 20 (2010), pp. 479–514.

[69] M. Barigozzi, G. Fagiolo, and D. Garlaschelli. "Multinetwork of international trade: A commodity-specific analysis". In: *Physical Review E* 81 (2010), p. 046104.

[70] M. Barigozzi, G. Fagiolo, and G. Mangioni. "Identifying the community structure of the international-trade multi-network". In: *Physica A* 390 (2011), pp. 2051–2066.

[71] J. McNerney, B. D. Fath, and G. Silverberg. "Network structure of inter-industry flows". In: *Physica A* 392 (2013), pp. 6427–6441.

[72] Y. Fan, S. Ren, H. Cai, and X. Cui. "The state's role and position in international trade: A complex network perspective". In: *Economic Modelling* 39 (2014), pp. 71–81.

[73] M. P. Timmer, E. Dietzenbacher, B. Los, R. Stehrer, and G. J. de Vries. "An illustrated user guide to the world input-output database: the case of global automotive production". In: *Review of International Economics* 23(3) (2015), pp. 575–605.

[74] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre. "Fast unfolding of communities in large networks". In: *Journal of Statistical Mechanics* 2008(10) (2008), P10008.

[75] Bureau of Labor Statistics. *Historical consumer price index for all urban consumers (CPI-U), U.S. city average, all items.* `https://www.bls.gov/cpi/historical-cpi-u-201708.pdf`. Table 24 (Accessed 20 September 2017).

[76] Z. K. Gao, X. W. Zhang, N. D. Jin, N. Marwan, and J. Kurths. "Multivariate recurrence network analysis for characterizing horizontal oil-water two-phase flow". In: *Physical Review E* 88 (2013), p. 032910.

[77] E. A. Bender and E. R. Canfield. "The asymptotic number of labeled graphs with given degree sequences". In: *Journal of Combinatorial Theory A* 24 (1978), pp. 296–307.

[78] International Monetary Fund. *World Economic Outlook, April 2009.* IMF Multimedia Services Division, 2009.

[79] D. Ricardo. *On the principles of political economy and taxation.* London: John Muarray, 1817.

[80] M. Cipollina and L. Salvatici. "Reciprocal trade agreements in gravity models: A meta-analysis". In: *Review of International Economics* 18(1) (2010), pp. 63–80.

[81] S. L. Baier and J. H. Bergstrand. "Do free trade agreements actually increase members' international trade?" In: *Journal of International Economics* 71(1) (2007), pp. 72–95.

[82] C. Carrère. "Revisiting the effects of regional trade agreements on trade flows with proper specification of the gravity model". In: *European Economic Review* 50(2) (2006), pp. 223–247.

[83] J. E. Anderson and Y. V. Yotov. "Terms of trade and global efficiency effects of free trade agreements, 1990–2002". In: *Journal of International Economics* 99 (2016), pp. 279–298.

[84] M. Dai, Y. V. Yotov, and T. Zylkin. "On the trade-diversion effects of free trade agreements". In: *Economics Letters* 122(2) (2014), pp. 321–325.

[85] C. Freund and E. Ornelas. "Regional Trade Agreements". In: *Annual Review of Economics* 2(1) (2010), pp. 139–166.

[86] E. W. Bond, R. Riezman, and C. Syropoulos. "A strategic and welfare theoretic analysis of free trade areas". In: *International Trade Agreements and Political Economy.* World Scientific Publishing, 2013. Chap. 8, pp. 101–127.

[87] R. Riezman. "Can bilateral trade agreements help induce free trade". In: *Canadian Journal of Economics* 32(3) (1999), pp. 751–766.

[88] V. Vicard. "Determinants of successful regional trade agreements". In: *Economics Letters* 111(3) (2011), pp. 188–190.

[89] P. Krugman. *Is bilateralism bad?* NBER Working Papers 2972. National Bureau of Economic Research, Inc, 1989.

[90] P. R. Krugman. "The move toward free trade zones". In: *Proceedings - Economic Policy Symposium - Jackson Hole* (1991), pp. 7–58.

[91] J. A. Frankel, E. Stein, and S.-J. Wei. "Regional trading arrangements: natural or supernatural". In: *American Economic Review* 86(2) (1996), pp. 52–56.

[92] A. Dür, L. Baccini, and M. Elsig. "The design of international trade agreements: Introducing a new dataset". In: *The Review of International Organizations* 9(3) (2014), pp. 353–375.

[93] J. Whalley. "Why do countries seek regional trade agreements?" In: *The Regionalization of the World Economy*. NBER Chapters. National Bureau of Economic Research, Inc, 1998, pp. 63–90.

[94] L. Baccini and J. Urpelainen. "International institutions and domestic politics: Can preferential trading agreements help leaders promote economic reform?" In: *The Journal of Politics* 76(1) (2014), pp. 195–214.

[95] R. E. Feinberg. "The Political Economy of United States' Free Trade Arrangements". In: *World Economy* 26(7) (2003), pp. 1019–1040.

[96] A. Antkiewicz and J. Whalley. "China's new regional trade agreements". In: *The World Economy* 28(10) (2005), pp. 1539–1557.

[97] K. Head and T. Mayer. "Chapter 3 - Gravity Equations: Workhorse,Toolkit, and Cookbook". In: *Handbook of International Economics*. Ed. by E. H. Gita Gopinath and K. Rogoff. Vol. 4. Elsevier, 2014, pp. 131 –195.

[98] A. Arenas, A. Cabrales, A. Díaz-Guilera, R. Guimerà, and F. Vega-Redondo. "Search and congestion in complex networks". In: *Statistical Mechanics of Complex Networks*. Ed. by R. Pastor-Satorras, M. Rubi, and A. Diaz-Guilera. Springer Berlin Heidelberg, 2003, pp. 175–194.

[99] M. E. J. Newman. "A measure of betweenness centrality based on random walks". In: *Social Networks* 27(1) (2005), pp. 39–54.

[100] V. M. R. Muggeo. "Estimating regression models with unknown break-points". In: *Statistics in Medicine* 22(19) (2003), pp. 3055–3071.

[101] H. Akaike. "A new look at the statistical model identification". In: *Selected Papers of Hirotugu Akaike*. Ed. by E. Parzen, K. Tanabe, and G. Kitagawa. Springer New York, 1998, pp. 215–222.

[102] World Trade Organization. *Regional Trade Agreements Information System*. `http://rtais.wto.org/UI/PublicMaintainRTAHome.aspx`. (Accessed 09 October 2017). 2017.

[103] D. Hummels, J. Ishii, and K.-M. Yi. "The nature and growth of vertical specialization in world trade". In: *Journal of International Economics* 54(1) (2001), pp. 75–96.

[104] R. Johnson and G. Noguera. "Proximity and production fragmentation". In: *American Economic Review* 102(3) (2012), pp. 407–4011.

[105] G. Grossman and E. Helpman. *Outsourcing in a global economy*. NBER Working Paper 8728. National Bureau of Economic Research, Inc, 2002.

[106] I. Arribas, F. Pérez, and E. Tortosa-Ausina. "Measuring globalization of international trade: Theory and evidence". In: *World Development* 37(1) (2009), 127–145.

[107] K. De Backer and N. Yamano. "The measurement of globalisation using international input-output tables". In: *OECD Science, Technology and Industry Working Papers* 3 (2007), pp. 23–24.

[108] A. M. Falzoni and L. Tajoliss. "International fragmentation of production and trade volatility: An analysis for the European countries". In: *Modern Economy* 6(3) (2015), pp. 358–369.

[109] W. Kohler. "Aspects of international fragmentation". In: *Review of International Economics* 12(5) (2004), pp. 793–816.

[110] R. Feenstran and G. Hanson. "Global production sharing and rising inequality: A survey of trade and wages". In: *Handbook of international trade*. Ed. by K. Choi and J. Harrigan. Oxford: Basil Blackwell, 2003, pp. 146–185.

[111] R. Baldwin and A. Venables. "Spiders and snakes: Offshoring and agglomeration in the global economy". In: *Journal of International Economics* 90(2) (2013), pp. 245–254.

[112] The World Bank. *Inflation, GDP deflator (annual %). World Bank national accounts data, and OECD national accounts data files*. `http://data.worldbank.org/indicator/NY.GDP.DEFL.KD.ZG`. (Accessed 04 October 2017).

[113] The World Bank. *Official exchange rate (LCU per US$, period average). International Monetary Fund, international financial statistics*. `http://data.worldbank.org/indicator/PA.NUS.FCRF`. (Accessed 04 October 2017).

[114] H. Görg and A. Hanley. "International outsourcing and productivity: evidence from the Irish electronics industry". In: *The North American Journal of Economics and Finance* 16(2) (2005), pp. 255–269.

[115] A. J.-Y. Kam. "International production networks and host country productivity: evidence from Malaysia". In: *Asian-Pacific Economic Literature* 27(1) (2013), pp. 127–146.

[116] L. A. N. Amaral, P. Gopikrishnan, V. Plerou, and H. E. Stanley. "A model for the growth dynamics of economic organizations". In: *Physica A* 299(1-2) (2001), pp. 127–136.

[117] R. Babbar, C. Metzig, I. Partalas, E. Gaussier, and M. Amini. "On power law distributions in large-scale Taxonomies". In: *ACM SIGKDD Explorations Newsletter* 16(1) (2014), pp. 47–56.

[118] M. F. Montes and V. V. Popov. *The Asian crisis turns global*. Singapore: ISEAS, 2000.

[119] A. Shama. "After the meltdown: a survey of international firms in Russia". In: *Business Horizons* 43(4) (2000), pp. 73–81.

[120] The National Bureau of Economic Research. *The NBER's Business Cycle Dating Committee*. `http://www.nber.org/cycles/recessions.html`. (Accessed 04 October 2017).

[121] J. J. Morris and A. Pervaiz. "Value relevance and the dot-com bubble of the 1990s". In: *The Quarterly Review of Economics and Finance* 52(2) (2012), pp. 243–255.

[122] G. A. Horn and U. Fritsche. "Argentina in crisis". In: *Economic Bulletin* 39(4) (2002), pp. 119–126.

[123] M. Palus. "From nonlinearity to causality: statistical testing and inference of physical mechanisms underlying complex dynamics". In: *Contemporary Physics* 48(6) (2007), pp. 307–348.

[124] D. G. Bonnett and T. A. Wright. "Sample size requirements for estimating pearson, kendall and spearman correlations". In: *Psychometrika* 65(1) (2000), pp. 23–28.

[125] S. Hempel, A. Koseska, and Z. Nikoloski. "Data-driven reconstruction of directed networks". In: *European Physical Journal B* 86(6) (2013), pp. 1–17.

[126] P. Y. Chen and P. M. Popovich. *Correlation: parametric and nonparametric measures*. Thousand Oaks: Sage, 2002.

[127] C. Spearman. "The proof and measurement of association between two things". In: *The American Journal of Psychology* 15 (1904), pp. 72–101.

[128] T. Schreiber and A. Schmitz. "Surrogate time series". In: *Physica D: Nonlinear Phenomena* 142(3-4) (2000), pp. 346–382.

[129] A. N. Kolmogorov. "Sulla determinazione empirica di una legge di distribuzione". In: *Giornale dell'Istituto Italiano degli Attuari* 4 (1933), pp. 83–91.

[130] J. Lin. "Divergence measures based on the Shannon entropy". In: *IEEE Transactions on Information Theory* 37(1) (1991), pp. 145–151.

[131] J. A. Schumpeter. *Capitalism, socialism and democracy*. Routledge, 2010.

[132] G. M. Angeletos and J. La'O. "Sentiments". In: *Econometrica* 81(2) (2013), pp. 739–779.

[133] H. Goto, H. Takayasu, and M. Takayasu. "Empirical analysis of firm-dynamics on Japanese interfirm trade network". In: *Proceedings of the International Conference on Social Modeling and Simulation, plus Econophysics Colloquium 2014* (2015).

[134] U. Alon. "Network motifs: theory and experimental approaches". In: *Nature Review Genetics* 8(6) (2007), pp. 450–461.

[135] P. V. Paulau, C. Feenders, and B. Blasius. "Motif analysis in directed ordered networks and applications to food webs". In: *Scientific Reports* 5 (2015), p. 11926.

[136] X. Zhang, S. Shao, H. E. Stanley, and S. Havlin. "Dynamic motifs in socio-economic networks". In: *Europhysics Letters* 108(5) (2014), p. 58001.

[137] R. Milo, N. Kashtan, S. Itzkovitz, M. E. J. Newman, and U. Alon. "On the uniform generation of random graphs with prescribed degree sequences". In: *arXiv:cond-mat/0312028* (2003).

[138] X. K. Xu, J. Zhang, and M. Small. "Changing motif distributions in complex networks by manipulating rich-club connections". In: *Physica A* 390(23-24) (2011), pp. 4621–4626.

[139] M. T. Angulo, Y. Y. Liu, and J. J. Slotine. "Network motifs emerge from interconnections that favour stability". In: *Nature Physics* 11(10) (2015), pp. 848–852.

[140] R. Albert and A.-L. Barabási. "Statistical mechanics of complex networks". In: *Reviews of Modern Physics* 74 (2002), pp. 47–97.

[141] Y. Fujiwara and H. Aoyama. "Large-scale structure of a nation-wide production network". In: *European Physical Journal B* 77(4) (2010), pp. 565–580.

[142] M. Takayasu, S. Sameshima, H. Watanabe, T. Ohnishi, H. Iyatomi, T. Iino, Y. Kobayashi, K. Kamehama, Y. Ikeda, H. Takayasu, and K. Watanabe. "Massive economics data analysis by econophysics methods-the case of companies network structure". In: *Annual Report of the Earth Simulator Center* (2007), p. 263.

[143] M. E. J. Newman. "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality". In: *Physical Review E* 64(1) (2001), p. 016132.

[144] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley. "Classes of behavior of small-world networks". In: *Proceedings of the National Academy of Sciences* 97 (2000), pp. 11149–11152.

[145] S. Redner. "How popular is your paper? An empirical study of the citation distribution". In: *European Physical Journal B* 4(2) (1998), pp. 131–134.

[146] M. Kitsak, L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse. "Identification of influential spreaders in complex networks". In: *Nature Physics* 6 (2010), pp. 888–893.

[147] O. Hein, M. Schwind, and W. König. "Scale-free networks". In: *Wirtschaftsinformatik* 48(4) (2006), pp. 267–275.

[148] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási. "The large-scale organization of metabolic networks". In: *Nature* 407 (2000), pp. 651–654.

[149] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, and J. Wiener. "Graph structure in the web". In: *Computer Networks* 33 (2000), pp. 309–320.

[150] R. Pastor-Satorras, A. Vázquez, and A. Vespignani. "Dynamical and correlation properties of the internet". In: *Physical Review Letters* 87 (2001), p. 258701.

[151] E. Ravasz and A.-L. Barabási. "Hierarchical organization in complex networks". In: *Physical Review E* 67(2) (2003), p. 026112.

[152] S. N. Dorogovtsev and J. F. F. Mendes. "Evolution of networks". In: *Advances in Physics* 51 (2002), pp. 1079–1187.

[153] T. Ohnishi, H. Takayasu, and M. Takayasu. "Network motifs in an inter-firm network". In: *Journal of Economic Interaction and Coordination* 5(2) (2010), pp. 171–180.

[154] D. Garlaschelli and M. I. Loffredo. "Patterns of link reciprocity in directed networks". In: *Physical Review Letters* 93(26) (2004), p. 268701.

[155] F. Leyvraz. "Scaling theory and exactly solved models in the kinetics of irreversible aggregation". In: *Physics Reports* 383(2-3) (2003), pp. 95–212.

[156] M. E. J. Newman. "Power laws, Pareto distributions and Zipf's law". In: *Contemporary Physics* 46(5) (2005), pp. 323–351.

[157] S. B. Moeller, F. Schlingemann, and R. M. Stulz. "Firm size and the gains from acquisitions". In: *Journal of Financial Economics* 73(2) (2004), pp. 201–228.

[158] D. Petmezas. "What drives acquisitions?: Market valuations and bidder performance". In: *Journal of Multinational Financial Management* 19(1) (2009), pp. 54–74.

[159] W. Miura, H. Takayasu, and M. Takayasu. "The origin of asymmetric behavior of money flow in the business firm network". In: *European Physical Journal Special Topics* 212 (2012), pp. 65–75.

[160] S. M. E. Sahraeian and B.-J. Yoon. "A Network Synthesis Model for Generating Protein Interaction Network Families". In: *PLOS ONE* 7.8 (2012), e41474.

Selbständigkeitserklärung

Ich erkläre, dass ich die Dissertation selbständig und nur unter Verwendung der von mir gemäß § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 126/2014 am 18.11.2014 angegebenen Hilfsmittel angefertigt habe.

Berlin, den 23. November 2017 Julian Maluck